DOCUMENT RESUME

ED 383 720 TM 023 141

AUTHOR	Henning, Grant
TITLE	A Study of the Effects of Variation of Short-Term Memory Load, Reading Response Length, and Processing Hierarchy on TOEFL Listening Comprehension Item Performance. Report 33.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-90-18
PUB DATE	Feb 91
NOTE	75p.
PUB TYPE	Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160) -- Statistical Data (110)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Adults; Auditory Stimuli; *Cognitive Processes; Educational Assessment; English (Second Language); *Listening Comprehension; *Reading; Responses; *Short Term Memory; Tables (Data); *Test Construction; Test Items; Test Length
IDENTIFIERS	Hierarchical Learning; Repetition (Language); *Test of English as a Foreign Language

ABSTRACT

Criticisms of the Test of English as a Foreign Language (TOEFL) have included speculation that the listening test places too much burden on short-term memory as compared with comprehension, that a knowledge of reading is required to respond successfully, and that many items appear to require mere recall and matching rather than higher-order processing skills. To address these criticisms, a study was designed with 120 learners for whom English was a Second Language (ESL learners) and 3 listening tests with 144 total items to examine test item functioning under conditions of stimulus repetition versus nonrepetition, variations of length of aural stimulus passage and of associated numbers of items, shorter versus longer reading response options, and higher versus lower levels of processing skills required. Overall, results suggest that tasks like those of the TOEFL listening test would benefit from shortening the response-option length, but that it would not be beneficial to repeat stimulus passages nor to increase the proportion of items that depend on comprehension of greater rather than lesser amounts of text. Nine tables present the analyses, and an appendix presents the test instruments. (Contains 17 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Marilynn Halpern

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED 383 720

TOEFL®

ENGLISH AS A FOREIGN LANGUAGE

Research Reports

REPORT 33
FEBRUARY 1991

A Study of the Effects of Variation of Short-Term Memory Load, Reading Response Length, and Processing Hierarchy on TOEFL Listening Comprehension Item Performance

Grant Henning

EDUCATIONAL TESTING SERVICE

TM 023 141

BEST COPY AVAILABLE

A Study of the Effects of Variation of Short-Term Memory Load, Reading Response Length, and Processing Hierarchy on TOEFL Listening Comprehension Item Performance

by

Grant Henning

Educational Testing Service

Princeton, NJ 08541

RR-90-18

Educational Testing Service is an Equal Opportunity/Affirmative Action Employer.

Copyright © 1991 by Educational Testing Service. All rights reserved.

Unauthorized reproduction in whole or in part is prohibited
without written permission from the publisher.

TOEFL, the TOEFL logo, ETS, and the ETS logo are trademarks of Educational Testing Service,
registered in the U.S.A. and in many other countries.
Educational Testing Service is a U.S.-registered trademark.

Acknowledgments

Seldom can research of this nature be conducted without the assistance of many competent individuals. The author is especially grateful for the help of Jean Turner of Santa Monica College and the University of California at Los Angeles, who scheduled and administered tests for about two-thirds of the experimental subjects. Access to the remaining subjects was graciously arranged by William Jex, Director of the American Language Institute of the New York University. Lisa Kaiser, also of New York University, assisted in the scheduling and administration of tests at that institution. Several research, program direction, and test development colleagues at Educational Testing Service routinely provided insightful comments on the proposal and the final report of the study. John H.A.L. de Jong of CITO, the National Institute for Educational Measurement in the Netherlands, provided helpful feedback at the proposal stage. Joanne Farr, of the Cognitive and Assessment Research Division of Educational Testing Service, provided valuable administrative assistance in the preparation of the tests and of the tables in this report. Participants at the 1989 San Antonio meeting of the Annual Language Testing Research Colloquium, especially Patricia Dunkel of Pennsylvania State University, provided feedback to work in progress that led to the incorporation of the particular measure of short-term memory employed in the study. Members of the TOEFL Research Committee encouraged and supported the study from the beginning.

The Test of English as a Foreign Language (TOEFL) was developed in 1963 by a National Council on the Testing of English as a Foreign Language, which was formed through the cooperative effort of more than thirty organizations, public and private, that were concerned with testing the English proficiency of nonnative speakers of the language applying for admission to institutions in the United States. In 1965, Educational Testing Service (ETS) and the College Board assumed joint responsibility for the program, and in 1973, a cooperative arrangement for the operation of the program was entered into by ETS, the College Board, and the Graduate Record Examinations (GRE) Board. The membership of the College Board is composed of schools, colleges, school systems, and educational associations; GRE Board members are associated with graduate education.

ETS administers the TOEFL program under the general direction of a Policy Council that was established by, and is affiliated with, the sponsoring organizations. Members of the Policy Council represent the College Board and the GRE Board and such institutions and agencies as graduate schools of business, junior and community colleges, nonprofit educational exchange agencies, and agencies of the United States government.

A continuing program of research related to the TOEFL test is carried out under the direction of the TOEFL Research Committee. Its six members include representatives of the Policy Council, the TOEFL Committee of Examiners, and distinguished English as a second language specialists from the academic community. Currently the Committee meets twice yearly to review and approve proposals for test-related research and to set guidelines for the entire scope of the TOEFL research program. Members of the Research Committee serve three-year terms at the invitation of the Policy Council; the chair of the committee serves on the Policy Council.

Because the studies are specific to the test and the testing program, most of the actual research is conducted by ETS staff rather than by outside researchers. However, many projects require the cooperation of other institutions, particularly those with programs in the teaching of English as a foreign or second language. Representatives of such programs who are interested in participating in or conducting TOEFL-related research are invited to contact the TOEFL program office. Local research may sometimes require access to TOEFL data. In such cases, the program may provide the data following approval by the Research Committee. All TOEFL research projects must undergo appropriate ETS review to ascertain that the confidentiality of data will be protected.

Current (1990-91) members of the TOEFL Research Committee are:

Patricia L. Carrell (Chair)	University of Akron
James Dean Brown	University of Hawaii
Patricia Dunkel	Pennsylvania State University
Fred Genesee	McGill University
Elliott Judd	University of Illinois at Chicago
Elizabeth C. Traugott	Stanford University

Abstract

This study was conceived in response to criticisms of the current TOEFL listening comprehension test-item format. Major areas of criticism have included speculation that listening as tested places too much burden on short-term memory as opposed to comprehension, that a knowledge of reading is required in order to respond successfully, and that many items appear to require mere recall and matching of details rather than higher-order processing skills. To address these criticisms in turn, a study was designed with 120 ESL learners and three listening tests (comprised of 144 total real and adapted TOEFL test items) to examine the characteristics of item functioning under conditions of stimulus repetition versus nonrepetition, variation of length of aural stimulus passage and of associated numbers of items, shorter versus longer reading response options, and higher versus lower level of processing skills required. Those item types and stimulus conditions that were found to associate with superior item functioning as indicated by estimates of item difficulty, item discriminability, internal consistency reliability, fit to a latent trait model, and convergent and discriminant validity were identified.

Results suggested that, while repetition of the stimulus passage predictably tended to reduce item difficulty when control was made for concomitant influences, there was no consistent effect of stimulus passage repetition on item discrimination, Rasch model fit, or discriminant validity across difficulty level. However, there was a tendency for items in the no-repetition condition to exhibit greater convergent and discriminant validity than items in the one-repetition condition.

Although passage length was confounded with numbers of items per passage and with comprehension hierarchy level, the test with passages of three-sentence length tended to be more reliable than the test with passages of two-sentence length, and the test with passages of two-sentence length tended to be more reliable than the test with passages of one-sentence length. Also, the test with the longest passages tended predictably to be slightly more difficult than the test with the shortest passages.

Item response-option length was significantly related to item difficulty and Rasch model fit in the direction that items with options that were shortened to about half current TOEFL response-option length tended to be easier and to exhibit better fit than items with current longer options. Also, items with shortened options showed greater convergent and discriminant validity across levels of difficulty than did items with unshortened options. And, there was a near-significant tendency for items with shortened options to exhibit better discrimination than items with unshortened options, when concomitant influences were controlled.

Comprehension hierarchy level of items, as defined by the length of passage required to respond correctly, was not significantly related to item difficulty except through a complex option-length-by-hierarchy-level interaction. However, hierarchy level was related to discrimination and Rasch model fit in the direction that items with lower level of processing (i.e., those that required comprehension of less stimulus text) showed better fit and discrimination than higher-level items after concomitant influences were removed. Also, greater convergent and discriminant validity across difficulty levels was exhibited by lower-level comprehension items than higher-level items.

It was concluded that tasks like those employed in TOEFL Listening Comprehension Section A would benefit from a shortening of current response-option length, but that it was not beneficial to repeat stimulus passages, nor was it desirable to increase the proportion of items that depended on comprehension of greater rather than lesser amounts of text.

TABLE OF CONTENTS

Introduction. 1
Problem 1
Purpose 3
Variables of the Study 3
Method 4
Sample 4
Instrumentation 4
Procedure 6
Analyses 6
Results 7
Test and Subtest Descriptive Statistics . . 7
Descriptive Statistics for Independent and Dependent Variables 8
Factorial Analyses of Variance and Covariance 9
Multitrait-Multimethod Validation . . . 11
Discussion and Conclusions 13
References 19
Tables 21
Appendix (Instruments) 31

List of Tables and Instruments

Table 1: Sample Description 21
Table 2: Design of Listening Comprehension Test Administration 22
Table 3: Descriptive Statistics for All Tests . . 23
Table 4: Descriptive Statistics for Item Variables . 24
Table 5: Correlations Among Selected Item Variables . 25
Table 6A: ANOVA with Rasch Difficulty as Dependent Variable 26
Table 6B: ANCOVA with Rasch Difficulty as Dependent Variable and with Sequence and Encounters as Concomitant Variables 26
Table 6C: Means and Standard Deviations of Rasch Item Difficulty Estimates at All Levels of Significant Effects 27
Table 7A: ANOVA with Overlap-Corrected Biserial Item-Total Individual Test Discrimination as Dependent Variable 28
Table 7B: ANCOVA with Overlap-Corrected Biserial Item-Total Individual Test Discrimination as Dependent Variable and with Sequence and Encounters as Concomitant Variables . . 28
Table 8A: ANOVA with Rasch Model Fit as Dependent Variable 29
Table 8B: ANCOVA with Rasch Model Fit as Dependent Variable and with Sequence and Encounters as Concomitant Variables 29
Table 9: Multitrait-Multimethod Validation Matrix with Two Levels Each of Repetition, Length, and Hierarchy in Place of Traits and High/Low Difficulty Items in Place of Methods . 30
Appendix (Instruments) 31
Questionnaire 33
Script - 1 35
Listening Comprehension - 1 37
Script - 2 42
Listening Comprehension - 2 45
Script - 3 50
Listening Comprehension - 3 53
TOEFL Reading Comprehension Test 58
Digital Memory Test 62

A. PROBLEM

The current TOEFL listening comprehension component has demonstrated highly satisfactory levels of internal-consistency reliability and criterion-related validity (Hale, Stansfield, & Duran, 1984; Pike, 1979). Nevertheless, at least three criticisms of this component have been expressed by some TOEFL users, including former members of the TOEFL Committee of Examiners and the TOEFL Research Committee who were requested to offer such criticisms. First, it has been alleged that the format used places too much load on short-term memory as opposed to comprehension. Secondly, it has been claimed that the use of a reading response format invalidates the test as a measure of listening comprehension only. Finally, it has been asserted that too many items require recall of minute details rather than higher-level processing strategies (Savignon, 1986; Stansfield, 1986). Other criticisms related to communicative focus and language authenticity have also been advanced (Bachman, 1986; Duran, Canale, Penfield, Stansfield, & Liskin-Gasparro, 1985). However, for the most part, these last concerns appear to have been discussed already by others (Larsen-Freeman, 1986; Oller, 1986) and are not included as foci in the present research.

With regard to the first criticism, related to memory versus comprehension becoming the focus of the test, it should be noted at the outset that some element of memory use would necessarily be present in any analysis of the listening comprehension construct. Unfortunately, as Carroll (1971), Devine (1978), and Larson, Backlund, Redmond, and Barbour (1978) have noted, there does not appear to be any scientific consensus about the exact nature of the listening comprehension construct or its components. Thus, there is no agreement on what portion of listening comprehension may be attributable to memory (whether short term or long term), or when that portion has been exceeded with any proposed listening comprehension task. What does appear possible to ascertain through a study of the kind presented here is whether variation of memory taxation in a listening comprehension task differentially and systematically affects item quality. Item quality can be determined operationally in terms of appropriate difficulty for the population of interest, higher rather than lower discriminability, internal consistency of item subgroupings, and greater convergent and discriminant validity of the particular item format referenced.

With regard to the criticism of the use of reading response options for listening comprehension items, several replies are possible, but each calls for empirical evidence. Carroll (1971)

conducted an extensive review of the literature of that time dealing with the comprehension of meaningful verbal discourse. Evidence gathered there suggests that use of a combined listening and reading presentation mode may be advantageous at some levels of learner proficiency, but may interfere at other levels. It may also be asserted that, because language skills are known to be highly intercorrelated in general (Oller, 1979), and, because at least the TOEFL listening passage stimulus and item stem or prompt are presented aurally rather than in writing, use of a multiple-choice reading response format for a listening comprehension task would not appreciably contaminate the validity of the component as a measure of listening comprehension. However, if some reading contamination were found to be present for some items through inspection of excessive correlations with an independent reading measure, it would become useful to discover ways of minimizing such contamination. This would be particularly true if the correlations of the listening items with the reading criterion were higher than the correlations of those same items with their own listening subscale total. One obvious way to minimize such potential contamination effects would be to reduce the length of the reading response task. The present study considers item quality as described above under two different levels of item response-option reading length, i.e., current TOEFL response-option length and an adaptation of current response-option length made by shortening response options to about half their current length.

With regard to the final criticism considered here, the one dealing with the cognitive processing hierarchy level addressed by the items, it has not always been easy for experts to reach consensus on exactly what constitutes higher and lower order of processing for any given set of comprehension items. Alderson (1986) found that experts could reach consensus on only one-third of a set of ESL reading comprehension items as to which items involved higher-order and which items involved lower-order cognitive processing. Even more disturbing for comprehension theorists has been his finding that, for those items for which consensus on classification was reached, the lower-order items systematically outperformed the higher-order items psychometrically. Due to this anticipated difficulty in achieving consensus on classification, for the present study items were classified by processing hierarchy in accordance with the breadth of stimulus passage information needed to be processed before the correct answer could be given. "Higher-order" items were those that required understanding of information across two or more sentences, while lower-order items could be answered correctly on the basis of understanding of information found in a word or phrase within a single sentence of stimulus discourse. Once again, it was thought possible to determine item quality with reference to the criteria listed above, this time with respect to items at differing levels of processing hierarchy.

In a related research study, Powers (1985) analyzed survey responses of 144 university professors from 28 institutions to determine, among other things, which listening comprehension tasks

were judged most appropriate for inclusion in a test of ESL listening comprehension. Of 23 general and specific tasks examined, responding to questions involving comprehension of numbers and numerical concepts, providing inferences and deductions, answering with recalled details, and condensing what is heard to outline form were the tasks most highly ranked by respondents. Although that study differed from the present one in that there was no investigation of item functioning in the former study, it is nevertheless interesting that there was no clear preference established for tasks involving higher-order over lower-order processing strategies.

B. PURPOSE

The present study was conducted to examine the effects of varying memory load through use of repetitive and nonrepetitive aural presentation procedures and through use of varying passage length formats. By varying repetition condition, passage length, and numbers of associated items, it was considered possible to investigate effects of these controlled variations on item difficulty, item discriminability, and format validity.

Additionally, the study was designed to consider the influences of varying length of reading task in the item response options. Two levels of reading length were examined (current TOEFL listening response option length and a systematically shortened version of the currently employed format). Again, the effects of varying option length were compared for measures of item difficulty, item discriminability, and format validity.

Of further interest was an investigation of the comparative performances of listening comprehension items at three levels of the processing hierarchy, from memory for details within single sentences, to memory for information presented across two sentences, to comprehension of information encountered across three passage sentences. Previous research in the measurement of reading comprehension has called attention to the difficulty of reliable classification of hierarchies of cognitive processing (Alderson, 1986). It was hoped that this strategy of classification according to extent of context upon which the item is based would help to overcome this classification difficulty in the case of listening comprehension assessment. Once again, comparisons of item difficulty, item discriminability, and format validity were made under each of the levels of processing hierarchy.

Specifically, the principal variables of interest in this study were:

(1) Repetition of stimulus passage. Two levels of repetition were considered: no repetition and one repetition.

(2) Passage length. Three levels of passage length were considered: passages of one-, two-, and three-sentence length,

containing approximately 10, 20, and 30 words, with one, two, or three associated test items, respectively.

(3) Reading response option length. Two levels of multiple-choice response-option length were considered: current TOEFL listening response-option length and a highly shortened version of the current TOEFL response-option length. The current TOEFL response-option length averaged 6.89 words with a standard deviation of 1.25 words for the 72 unshortened items in the study. The shortened option length averaged 3.34 words with a standard deviation of 0.76 words for the 72 shortened items of the study. Thus, on average, the shortened options were slightly less than half the length of the unshortened options. Additionally, it should be noted that stems consisting of from one to three words were added to many of the shortened items to facilitate reduction of overall length. The stems averaged 1.79 words in length with a standard deviation of 0.91 words across the 72 shortened items.

(4) Processing hierarchy. Items were designed to measure three levels of comprehension: comprehension of discrete details within single sentences, comprehension of information presented across two sentences, and comprehension of information presented across three sentences. Thus, processing hierarchy was defined operationally in terms of the comparative length of the stimulus passage required to be processed in order to obtain the answer to the item.

C. METHOD

1. Sample

A sample of 120 subjects was identified from among the English-as-a-second-language (ESL) students at three U.S. schools (Santa Monica Community College, UCLA Extension, and New York University American Language Institute). Subjects varied widely in language proficiency, language background, time of residence in an English-speaking country, and time of formal English language study. (See Table 1 for a summary of subject characteristics.) All subjects volunteered to participate in consideration of the test-taking practice opportunity, the award of TOEFL practice materials, or nominal equivalent monetary compensation.

2. Instrumentation

The following instruments were designed or adapted for the study:

(a) A brief, one-page demographic questionnaire requesting information about native language background, length of residence in an English-speaking country, and length of English language study.

(See the Appendix for a copy of this questionnaire.)

(b) A listening comprehension test with 48 one-sentence stimulus passages and 48 related items (1 item per passage). These passages were varied so that 24 were repeated once and 24 were not repeated. Of the 24 repeated passages (24 items) and again of the 24 nonrepeated passages (24 items), 12 associated items exhibited shortened reading response-option format and 12 associated items exhibited current unshortened reading response-option format. All items were copied or adapted from prior, disclosed TOEFL forms. (See the Appendix for a copy of this test, labeled "Listening Comprehension - 1.")

(c) A listening comprehension test with 24 two-sentence stimulus passages and 48 related items (2 items per passage). These passages were varied so that 12 were repeated once and 12 were not repeated. Of the 12 repeated passages (24 items) and again of the 12 non-repeated passages (24-items), 12 associated items exhibited shortened reading response option format and 12 associated items exhibited current unshortened reading response option format. Distributed evenly and systematically throughout the test were items representing two levels of processing hierarchy (24 items for level one and 24 items for level two as described above--one item of each level for each passage). All items were copied or adapted from prior disclosed TOEFL forms. (See the Appendix for a copy of this test, labeled "Listening Comprehension - 2.")

(d) A listening comprehension test with 16 three-sentence stimulus passages and 48 related items (3 items per passage). These passages were varied so that 8 were repeated once and 8 were not repeated. Of the 8 repeated passages (24 items) and again of the 8 nonrepeated passages (24 items), 12 associated items exhibited shortened reading response-option format and 12 associated items exhibited current unshortened reading response-option format. Distributed evenly and systematically throughout the test were items representing three levels of processing hierarchy (16 items for each of levels one, two and three as described above--1 item of each level for each passage). All items were copied or adapted from prior, disclosed TOEFL forms. (See the Appendix for a copy of this test, labeled "Listening Comprehension - 3.")

In construction of the three 48-item listening comprehension tests described in b, c, and d above, use was made of the items in only part A of the listening comprehension components of the disclosed forms from the August 1985, July 1986, and November 1987 administrations of the TOEFL test.

(e) A disclosed TOEFL reading comprehension component test was administered to all subjects to provide a concomitant measure of reading ability. (See the Appendix for a copy of this test.)

(f) A 15-item digital memory test was administered to all subjects to provide a concomitant measure of short-term memory. (See

the Appendix for a copy of this test.)

3. Procedure

All 120 subjects responded to all tests and questionnaires. To ensure that each listening test was encountered in every sequence by the same number of subjects to control for practice and sequence effects, the subjects were sequentially assigned to three permutations of sequence order (approximately 40 subjects per permutation). Thus, for person group one, listening comprehension tests were administered in the sequence 1, 2, 3. For person group two, listening comprehension tests were administered in the sequence 2, 3, 1. For person group three, listening comprehension tests were administered in the sequence 3, 1, 2. To control for within-test sequence effects, items were coded numerically by sequence and the resulting sequence variable was employed as a concomitant variable in the study after homogeneity of regression assumptions were shown to be satisfied. Also to minimize practice and sequence effects, feedback on item success or failure was not given at any time during test administration. Total testing time for all tests did not exceed two hours per subject. Balanced subsets of items and their associated stimulus sentences appeared in more than one test form. To control for any possible multiple-encounter effect, items were also coded numerically in accordance with the number of encounters across tests. The resulting encounters variable was employed as a concomitant variable after homogeneity of regression assumptions were shown to be satisfied. See Table 2 for a more thorough representation of the experimental design.

4. Analyses

Means, standard deviations, and internal consistency reliabilities were calculated for each test and subtest variation. Sample demographic information was also tallied.

Rasch model item difficulty and fit estimates and both biserial and point biserial item-total score discriminability indices were computed for every item under every response condition. Biserial and point-biserial correlations of every listening comprehension item were computed with the digital memory and TOEFL reading test scores. Mean and standard deviation item difficulty, discriminability, Rasch model fit, item-TOEFL reading correlation, and item-digital memory correlation were computed for all listening comprehension items. Use of Rasch model item difficulty estimates was preferred over traditional proportion correct (p) values because the former estimates provided a small-sample logarithmic transformation to an equal-interval scale (Wright & Stone, 1979). Biserial item-total, biserial item-TOEFL reading, and biserial item-digital memory correlations were preferred over their point-biserial counterparts because of the assumptions of normality of distribution that were believed tenable

for these data. Correlational computations employed correction for part-whole overlap and Fisher Z transformation as needed. The Rasch model fit statistic "infit" was employed as an item construct validity criterion (Wright & Linacre, 1984). Essentially, this fit statistic reports the degree of improbability of the pattern of responses to any item, given the pattern of responses of the same persons to all other items.

Factorial analyses of variance were calculated sequentially using Rasch item difficulty and fit estimates and item discrimination indices as dependent variables. Following appropriate tests of the assumption of homogeneity of regression coefficients, analyses of covariance were conducted using item sequence and item encounters as concomitant variables to test main effects and interaction effects with potentially contaminating influences removed.

Construct validity of the various item formats was assessed in two different ways. First, since Rasch model fit estimates provide an indication of the fit or response validity of the items to the expectations of the model, the forementioned ANOVA and ANCOVA procedures using fit as a dependent variable served to indicate the comparative validity of items under the various response conditions. Secondly, use was made of a procedure analogous to multitrait-multimethod validation procedure (Campbell & Fiske, 1959), with formats serving as traits and high-difficulty/low-difficulty item splits within formats serving to define methods. By this procedure each item format variation was examined for convergent and discriminant validity across levels of item difficulty. The comparative validities of format variations were ascertainable as comparative magnitudes of matrix diagonal coefficients.

D. RESULTS

1. Test and Subtest Descriptive Statistics

Descriptive statistics for all tests and subtests are provided in Table 3. Note that estimates of internal consistency reliability (alpha) are provided for every test and item-subtest combination. Note also that, since reliability is a partial function of the number of items in a test, the final column of the table provides Spearman-Brown adjusted estimates to hold the number of items constant at 50 for all tests and subtests.

These gross statistics across subtests reveal few significant differences that may be attached to particular item formats. The reported total test means reveal a predictable but slight tendency for tests with the longest passages (e.g., Listening Comprehension 3) to be most difficult and tests with the shortest passages (e.g., Listening Comprehension 1) to be least difficult. Repetition of the

stimulus passages showed no consistent difference from nonrepetition, whether in terms of test difficulty or of test reliability. With the exception of the first test (Listening Comprehension 1), there was a tendency for subtests with short reading response option items to be both easier and more reliable than subtests with longer reading response options. Subtests with items at levels of cognitive processing hierarchy as defined show a distinct tendency such that subtests with lower-order items tended to be both easier and more reliable than subtests with higher-order items. The shortest-passage test, Listening Comprehension 1, was less reliable than the second-shortest-passage test, Listening Comprehension 2 (.835 versus .871), and the second-shortest-passage test, Listening Comprehension 2, was less reliable than the longest-passage test, Listening Comprehension 3 (.871 versus .890). However, it should be noted that passage length was confounded with numbers of items per passage and with comprehension hierarchy levels of associated items.

2. Descriptive Statistics for Independent and Dependent Variables

Table 4 reports means, standard deviations, standard errors and ranges for the item variables employed in the study. Note also that every item was classified according to level of repetition (1=one repetition, 2=no repetition), level of response option length (1=shortened, 2=current length), and level of cognitive processing hierarchy (1=comprehension of information from a word or phrase within one sentence, 2=comprehension of information across two sentences, and 3=comprehension of information across three sentences of the stimulus passage). Table 4 reports results for two different difficulty statistics, six different discrimination statistics, and three item validity indicators. For reasons already given, some of these statistics were more appropriate than others for use in the subsequent analyses. Only those statistics deemed appropriate were subsequently employed in analyses.

The statistics reported for TOEFL reading and digital memory consist of the means, standard deviations, and ranges of biserial correlations computed between individual item scores and reading and recall test scores. Similarly, the statistics reported for discrimination consist of means, standard deviations, and ranges of both biserial and point biserial correlations between individual item scores and listening comprehension test total scores. In all analyses involving computation with correlation coefficients, use was made of Fisher Z transformations to correct for scaling inadequacies of correlation coefficients.

The Rasch model difficulty and fit data were estimated with Microscale Version 1.20 (Wright & Linacre, 1984). As a feature of that program, the mean of item difficulty statistics is arbitrarily set at zero. The particular fit statistic chosen was the Rasch model "infit" estimate also provided by that program. This is a sensitive

index of the degree of departure of individual item responses from model expectation. High positive fit statistics are usually interpreted as reflective of model misfit, while high negative fit statistics are said to represent overfit to the model (Wright & Stone, 1979).

Intercorrelations among a pertinent subset of these variables are reported in Table 5. The coefficients reported in Table 5 indicate that there was a weak positive relationship (0.174) between item difficulty (DIF) and the extent to which the item required processing of longer versus shorter segments of the listening stimulus passage (HIR). Item correlation with the digital memory (MEM) test showed a comparatively strong positive relationship to item correlation with the TOEFL reading test (RDG) (0.397), to item-total biserial discriminability (ITB) (0.381), and to Rasch model item infit (FIT) (-0.361). (Note that a negative correlation with the fit statistic reflects a positive relationship to model fit.) A similar pattern of correlations was observed for item correlations with reading test scores (RDG) as was observed for item correlations with recall test scores (RCL) discussed in this paragraph. Separate analyses indicated that only 15 of 144 items showed higher correlation with TOEFL reading than with their respective corrected domain totals, suggesting that 90 percent of all listening comprehension items could not be said to be contaminated by reading effects. Of those 15 deviant items, no clear frequency pattern was present for items of any one experimental condition over any other experimental condition. Similarly, only 11 of 144 items showed higher correlation with digital memory than with their respective corrected domain totals, suggesting that 92 percent of all listening comprehension items could not be said to be contaminated by memory effects. Interestingly, 10 of those 11 deviant items were of the lowest comprehension hierarchy level, implying that recall of discrete information within a single sentence was more taxing on memory than was recall of information across two or three sentences. No other patterns emerged for these items. Discrimination was related to difficulty (-0.185), Rasch fit (-0.763), TOEFL reading (0.433), and digital memory (0.381). In general, the nonsignificant correlations reported among repetition, option length, and hierarchy level with many of the other relevant item variables contrasted sharply with the results of the ANOVA, ANCOVA, and multitrait-multimethod type analyses that follow. These differences may be attributed to the effects of removal of interaction effects in the partitioning of variance, to the effects of removal of contributions of concomitant variables (in the case of ANOVA and ANCOVA), or to the effects of grouping items more directly within response conditions (in the case of the multitrait-multimethod type analysis).

3. Factorial Analyses of Variance and Covariance

(a) Effects on Difficulty

The effects of levels of repetition, option response reading

length, and processing hierarchy on Rasch model item difficulty estimates are reported in Tables 6A, 6B, and 6C. The analysis of variance reported in Table 6A indicates a significant effect of option length on item difficulty ($p = 0.018$); however, this generalization must be qualified by the finding of a significant interaction effect between length and hierarchy ($p = 0.015$). There was also a near-significant tendency toward an effect of repetition on item difficulty ($p = 0.066$). This tendency became even more salient in the analysis of covariance reported in Table 6B after concomitant influences of item sequence and item encounters were controlled. The means reported in Table 6C indicate the direction of the important effects noted in Table 6A. Changing from one repetition of the stimulus passage to no repetitions tended to increase item difficulty. Changing from shortened option length to current longer option length tended to increase item difficulty. Lower-order within-sentence processing items tended to be easier than higher-order across-sentence processing items, but the highest-order three-sentence items were not more difficult than the second-order two-sentence items. The significant length by hierarchy interaction effect was of a sort that, while difficulty did tend to increase with increase in response option length overall, at the lowest level of the processing hierarchy, items with shortened option length appeared more difficult than items with current unshortened length. However, at the second and third levels of the processing hierarchy, shortened option length was more strongly associated with lower item difficulty than was current longer option length.

Table 6B reports the results of analysis of covariance using the item sequence and item encounters variables as the two concomitant variables. These variables satisfied the ANCOVA assumption of homogeneity of regression slopes. As Table 6B indicates, while use of these concomitant variables in the analysis increased the power of testing, it did not alter the pattern of significance of the effects or the interpretation of outcomes. Multiple correlation coefficients accompanying each ANOVA and ANCOVA table provide some indication of the overall size of effects.

(b) Effects on Discrimination

Tables 7A and 7B report the effects of stimulus repetition, option length, and processing hierarchy on item discriminability as computed by item-total biserial correlation. The slight tendency for option length to affect item discriminability, noted in the ANOVA of Table 7A ($p = 0.129$), became more salient in the ANCOVA of Table 7B ($p = 0.075$), where the contributions of item sequence and item encounters are controlled in the same manner as was reported in Table 6B. Again, the ANCOVA assumption of homogeneity of regression coefficients was satisfied for the present analysis. Unlike the case with effects on difficulty reported earlier, there were no significant interaction effects in Tables 7A or 7B. The direction of the tendency of option length to affect discriminability was such that shorter option length was associated with greater discriminability than was current longer

option length. The ANCOVA of Table 7b reports a significant effect of hierarchy level on discrimination ($p = 0.036$). This effect was in the direction that items of lower levels of the comprehension hierarchy tended to show greater discrimination than did items of higher levels. These same results were replicated for these data when point-biserial correlations were used instead of biserial item-total score correlations to reflect item discriminability. It is important to note again that in all of these analyses Fisher Z transformations were used to enable more accurate computation with correlation coefficients.

(c) Effects on Model Fit

A final item quality criterion used in the study of impact of stimulus repetition, response-option length, and processing hierarchy was the Rasch model infit statistic generated by the software program Microscale 1.20. Since this statistic is sensitive to violations of unidimensionality constraints as would occur if respondents differentially guessed answers to some items or if some items tended to measure unintended constructs, the following analyses tend to reflect the comparative construct or response validity of the items under various response conditions. Tables 8A and 8B report the effects of stimulus repetition, option length, and processing hierarchy on Rasch model infit. Table 8A provides ANOVA information indicating significant option length ($p = 0.040$) and processing hierarchy ($p = 0.045$) effects on model fit. The ANCOVA of Table 8B indicates that no significant differences in the pattern of effects were observed after influence of the concomitant variables was controlled. The direction of these effects was such that items with shorter response options tended to provide better fit to the predictions of the model than did items with current longer response options, and items at levels one and two of the processing hierarchy tended to provide better fit to model expectations than did items at level three. Overall, items of lower-order processing hierarchy showed better construct validity than items of higher-order processing hierarchy as defined here and as judged in terms of impact on Rasch model infit.

4. Multitrait-Multimethod Validation

The Campbell and Fiske (1959) multitrait-multimethod validation procedure provides a set of criteria for establishing construct validity of proposed traits through inspection of an appropriate trait-by-method correlation matrix. An adaptation of this procedure was made for the present analysis in order to determine whether patterns of item responses under the various repetition, option length, and comprehension hierarchy levels would be stable across levels of item difficulty. This is a clear indication of the construct validity of tests comprised of items of the various format types (e.g., with passage repetition or no repetition, with shorter or longer response options, and with lower or higher levels of the

comprehension hierarchy). By this procedure, item conditions, here analogous to traits, would be judged to exhibit monotrait-heteromethod convergent validity if the correlations of test scores for the same traits across different levels of difficulty, here analogous to methods, were significantly greater than zero. As an additional step in the procedure, if the convergent validity coefficients considered in step one were also found to exceed in magnitude all adjacent heterotrait-monomethod coefficients, the traits associated with the convergent validity coefficients could be said to exhibit heterotrait-monomethod discriminant validity. And finally, if the convergent validity coefficients were found to exceed all adjacent heterotrait-heteromethod coefficients, the traits associated with the convergent validity coefficients could be said to exhibit heterotrait-heteromethod discriminant validity.

Table 9 reports the multitrait-multimethod validation matrix that was derived from intercorrelations of scores from twelve 12-item tests assembled purposefully from the items of the experimental tests in the present study. These twelve 12-item tests were formed by grouping separately high-difficulty items and low-difficulty items within two levels of stimulus repetition, option length, and processing hierarchy. For purposes of the analysis, high and low difficulty item groupings were considered analogous to methods in each trait comparision. Thus, construct validity in this study would reflect stability across the difficulty continuum of the item characteristic that is being considered. Note that the underscored coefficients in the diagonal of the matrix comprise the convergent validity coefficients, and all of these coefficients significantly exceed zero, so all traits show convergent validity by this lenient criterion. However, only the tests with items of shortened option length (LEN1, $r = 0.735$) and the tests with items of lowest-order processing hierarchy (HIR1, $r = 0.723$) exhibited discriminant validity in all required comparisons. While tests prepared from items of neither stimulus repetition condition were completely successful in terms of every discriminant validity comparison, the no-repetition condition (REP2, $r = 0.547$) showed greater convergent and discriminant validity than the one-repetition condition (REP1, $r = 0.437$).

Results of this analysis support the use of nonrepeated listening stimuli over repeated listening stimuli. Nonrepetition of listening stimuli is the current procedure with TOEFL listening comprehension testing. Furthermore, the analysis provides further support for use of item format that is shortened in option response length from the currently used option length. And finally, the analysis does not provide evidence in support of item format that requires higher- rather than lower-order cognitive processing as determined by the comparative length of the stimulus passage that must be processed to respond to the item.

E. DISCUSSION AND CONCLUSIONS

This study was proposed to compare TOEFL listening comprehension item quality under a variety of conditions of stimulus repetition, response-option reading length, and cognitive processing hierarchy. Results support several conclusions relevant to TOEFL listening comprehension test item development.

1. Repetition and Length of Stimulus Passage and Memory Effects

One concern expressed by critics of the current TOEFL listening format is that it places too much burden on short-term memory as opposed to tapping comprehension. The present study attempted to investigate this concern in several ways. First of all, item performance was examined under two repetition-of-stimulus conditions (i.e., one repetition and no repetition). The rationale for this procedure was that it was believed that repetition of the stimulus passage would lessen the burden on memory and permit a test of the effects of such a reduced burden on the performance of the associated items. While repetition also increases the opportunity to comprehend, it was thought that repetition would also reinforce memory for information that was comprehended on the first exposure. Results suggested that, while there was a predictable trend for items in the stimulus-repetition condition to be easier than items in the nonrepetition-of-stimulus condition (Table 6B, $p = 0.052$), there was no evidence that repetition of stimulus had any positive effect on item discrimination (Tables 7A and 7B), item response validity as indicated by fit to a latent-trait model (Tables 8A and 8B), or format construct validity as indicated by a procedure analogous to the Campbell and Fiske (1959) multitrait-multimethod validation procedure (Table 9).

To investigate the effects of length of stimulus passage on the quality of item performance, stimulus passages were constructed with lengths varying from one to three sentences. Here it was thought that length of stimulus passage could also provide a measure of burden on memory. Again, while there was a predictable tendency for tests composed of items associated with one-sentence stimulus passages to be easier than tests composed of items associated with two-sentence stimulus passages and for tests composed of items associated with two-sentence stimulus passages to be easier than tests composed of items associated with three-sentence stimulus passages (Table 3), test reliability tended to increase with increase in length of the stimulus passage (Table 3). Internal consistency reliability estimates for tests of 50-item length varied according to length of stimulus passage as follows: one-sentence passages, 0.841; two-sentence passages, 0.875; and three-sentence passages, 0.894. Since estimates of internal consistency can be shown to be positively related to item discriminability and increased potential for empirical validity, there is no evidence in the present results to suggest that any additional burden on memory associated with either stimulus passage length or nonrepetition of stimulus passage will negatively affect item quality or task validity.

A final consideration relevant to the issue of memory load in the assessment of listening comprehension involved the use of an independent measure of short-term memory as one of the tests in the study. Examinee scores on the memory test were correlated with the 120-case binary response vectors for each listening comprehension item. The resulting correlation coefficients provided one indication of the extent to which success with any given listening item was related to short-term memory. For the 144 listening comprehension items in this study, these memory-dependedness correlations were again correlated with such item characteristics as whether the item was associated with a repeated or a non-repeated stimulus passage, whether the written response options for items were of the shortened or unshortened variety, and whether the level of comprehension hierarchy was one, two, or three, as defined, for any given item (Table 5). Results suggested that, while memory dependedness was an important item characteristic as indicated by its significant correlations with estimates of item discriminability and model fit, there was no significant relationship (whether attenuated or disattenuated) between memory dependedness of item success and level of repetition, option length, or comprehension hierarchy. Only 11 of the 144 items were found for which the correlation with the digital memory test score exceeded the corrected correlation with subtest total score. Thus, 92 percent of the listening comprehension items showed greater relation to a measure of comprehension than to a measure of memory. Interestingly, 10 of the 11 deviant items were lowest-comprehension-level items, suggesting that correctly responding to items requiring comprehension of information within a single sentence was more taxing on memory than was correctly responding to items requiring comprehension of information across two or three sentences.

These results, taken separately and in combination, provide no support for the hypothesis that the current item formats in the TOEFL listening comprehension component overly tax short-term memory to the detriment of appropriate assessment of listening comprehension. While reduction of memory load of listening comprehension items would tend to result in easier items and higher test scores, such reduction of memory load would likely also be associated with reduction in both item discriminability and fit to a latent-trait model, and the reliability and validity of resulting tests would thereby be decreased.

2. Response Option Length and Reading Effects

Another concern of some critics of the present TOEFL listening item formats is related to the reliance on written-response options to assess the ability to comprehend spoken discourse. These critics would maintain that use of a reading task in the assessment of listening comprehension serves to confound the assessed construct of listening comprehension with that of reading comprehension. Thus, it is alleged that the listening component is not so valid as it would be if the response options were presented aurally rather than in writing.

It should be acknowledged that there may be program-operational constraints associated with time available for testing that dictate some such format decisions within large-scale testing programs such as the TOEFL program. Furthermore, since, for TOEFL, primary use is made of examination total scores rather than of component scores for decision-making purposes, use of more nearly integrative tasks within components would in no way compromise the validity of TOEFL total scores for intended uses. Nevertheless, it is an appropriate research question to determine whether the nature of the response task in the listening comprehension component of TOEFL could be altered in any way to improve the validity of that component.

To investigate the concern raised here, use was made of two levels of reading response-option length. 72 items were employed that used the current TOEFL response-option length (an average of 6.89 words per option), and 72 items were employed using an edited and highly shortened response-option length (an average of 3.34 words per option). The rationale for this procedure was that reduction of the usual reading task by about one half would enable a partial test of the value of such minimization of reading within the listening comprehension component. All of the items used were either actual, disclosed TOEFL listening items or were adapted from such items.

The results of several analyses (all but the correlational analysis of Table 5) suggested predictably that the items with shortened option length were easier than items with unshortened option length (Tables 3, 6A, 6B, and 6C), although the ANOVA and ANCOVA results were qualified somewhat by the finding of a significant interaction between length and processing hierarchy such that there was a tendency for items with shortened options to diminish in difficulty at the highest level of processing hierarchy (i.e., at the point where the task required synthesis over the greatest amount of passage content). Items with current, unshortened option length conversely tended to increase in difficulty with the increase in level of processing (Tables 6A, 6B, and 6C).

There was a nonsignificant tendency ($p = 0.075$) for items with shortened options to demonstrate greater discriminability than items with unshortened options when sequence and encounter effects were controlled through analysis of covariance (Table 7B). There was a significant tendency ($p = 0.040$, 0.033) for items with shortened option length to demonstrate greater response validity than items with unshortened option length as indicated by effects on fit to a latent-trait model (Tables 8A and 8B). Also, results of an analysis analogous to multitrait-multimethod analysis (Table 9) indicated that items with shortened option length, unlike items with unshortened option length, demonstrated discriminant validity in all required comparisons.

Although the present study was not designed to address fully the question of use of aural response options versus current written response options, results of several of the present analyses do

suggest that a reduction of current written response-option length by about one-half does lead to improved item discriminability and greater format validity. It is useful to observe that shortening the response options in the TOEFL listening component was one of the earlier recommendations offered independently by Savignon (1986).

3. Level of Processing and Comprehension Hierarchy Effects.

A final concern of interest here regards the criticism sometimes made of the listening comprehension component of the TOEFL test that too much reliance is placed on item types that tap comprehension at the lowest level (i.e., bottom-up comprehension or memory for discrete details in the stimulus passage) as opposed to higher levels of comprehension involving top-down strategies, such as inferencing and synthesizing processes. Related research in the area of reading comprehension has reported difficulty in obtaining expert agreement on what levels in the comprehension hierarchy are addressed by particular comprehension items (Alderson, 1986). To avoid this classification problem, in the present study distinctions among levels of processing were made on the basis of the amount of stimulus passage required to be processed in order to respond correctly to the test item. Items were designed accordingly at three levels of the comprehension hierarchy--that is, items requiring information successively from one, two, or three sentences of the stimulus passage in order to permit correct responding. Analyses were made of the comparative performances of the three item types.

Results suggested that, while there was a slight tendency for subtests comprised of lower-order comprehension items to exhibit higher mean scores and higher reliability estimates than did subtests with higher-order items (Table 3), there was no consistent effect on item difficulty associated with level of comprehension processing hierarchy across the 144 items and 120 persons in the present study (Tables 6A, 6B, and 6C). The one possible exception involves a significant interaction effect between option length and processing level that was discussed earlier.

There was a significant effect on item discriminability associated with comprehension processing level as defined (Table 7B). This effect was such that items representing lower levels of the comprehension hierarchy tended to discriminate better than items representing higher levels. Also, there was a significant effect of processing level on Rasch model fit detected in the ANOVA reported in Table 8A (such that lower-order items demonstrated better fit to the expectations of the model and, thus, greater response validity than did higher-order items). This effect persevered when sequence and encounter scores were used as concomitant variables in the ANCOVA reported in Table 8B. It is also possible, since there were more lower-order than higher-order items in the study, that fit to the expectations of the model would entail greater conformity to the response characteristics of the lower-order items and would thus bias the fit statistic in favor of lower-order items. The multitrait-multimethod analysis (Table 9) suggested that items of lower-order

processing level, unlike items of higher-order processing level, exhibited discriminant validity in all required comparisons.

These results do not support the view that a concerted effort should be made to ensure that a preponderance of TOEFL listening comprehension items be designed at the higher levels of the comprehension processing hierarchy. To the contrary, it would appear from the multitrait-multimethod analysis conducted that increased reliance on so-called lower-order items as defined for the present study may result in a commensurate increase in construct validity of the tests. It must be cautioned, however, that the practice of defining levels of processing for listening comprehension assessment by means of measures of the amount of discourse needed to be processed in order to respond correctly is not the only way of defining the comprehension hierarchy. Nevertheless, results of application of the present procedure appeared to underscore the psychometric value of lower-order comprehension items in the same way that Alderson's (1986) study supported their use in the assessment of reading comprehension.

It should also be noted here that the current TOEFL listening comprehension component includes a variety of item types, not all of which were systematically considered in the present study. The study was further limited by its primary focus on three identified concerns related to listening comprehension item format. Of the three major concerns investigated--memory load, reading response, and comprehension hierarchy--results suggested that the one concern with greatest merit toward the implementation of possible improvements in the format of TOEFL listening comprehension items is the concern related to the length of the reading response options. Several of the analyses indicated that reducing the length of the reading response options in a listening comprehension test such as TOEFL Listening Comprehension Section A by as much as one-half the current length could result in enhanced item and test quality in terms of a variety of established psychometric criteria. It is therefore recommended that appropriate consideration be given to the reduction of response-option length in the development of future versions of the listening component of the TOEFL test.

References

Alderson, J. C. (March, 1986). Processing hierarchies for reading comprehension items. A paper presented at the annual TESOL Convention, Anaheim, California.

Bachman, L. F. (1986). The Test of English as a Foreign Language as a measure of communicative competence. In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report 21). Princeton, N.J.: Educational Testing Service.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Carroll, J. B. (1971). Learning from verbal discourse in educational media: a review of the literature (ETS Research Bulletin RB-71-61). Princeton, NJ: Educational Testing Service.

Devine, T. G. (1978). Listening: What do we know after fifty years of research and theorizing? Journal of Reading, 21, 296-304.

Duran, R. P., Canale, M., Penfield, J., Stansfield, C. W., & Liskin-Gasparro, J. E. (1985). TOEFL from a communicative viewpoint on language proficiency: A working paper (TOEFL Research Report 17; ETS Research Report 85-8). Princeton, NJ: Educational Testing Service.

Hale, G. A., Stansfield, C. W., & Duran, R. P. (1984). Summaries of studies involving the Test of English as a Foreign Language, 1963-1982 (TOEFL Research Report 16). Princeton, NJ: Educational Testing Service.

Larsen-Freeman, D. (1986). A response to Sandra Savignon's "The meaning of communicative competence in relation to the TOEFL program." In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report 21). Princeton, NJ: Educational Testing Service.

Larson, C., Backlund, P., Redmond, M., & Barbour, A. (1978) Assessing functional communication. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills; Falls Church, VA: Speech Communication Association.

Oller, J. W., Jr. (1979). Language tests at school. London: Longman.

Oller, J. W., Jr. (1986). Communication theory and testing: What and how. In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report 21). Princeton, NJ: Educational Testing Service.

Pike, L. W. (1979). An evaluation of alternative item formats for testing English as a foreign language (TOEFL Research Report 2; ETS Research Report 79-6). Princeton, NJ: Educational Testing Service.

Powers, D. E. (1985). A survey of academic demands related to listening skills (TOEFL Research Report 20; ETS Research Report 85-48). Princeton, NJ: Educational Testing Service.

Savignon, S. J. (1986). The meaning of communicative competence in relation to the TOEFL program. In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report 21). Princeton, NJ: Educational Testing Service.

Stansfield, C. W. (Ed.) (1986). Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report 21). Princeton, NJ: Educational Testing Service.

Wright, B., & Linacre, J. M. (1984). Microscale manual. Black Rock, CT: Mediax Interactive Technologies.

Wright, B. D., & Stone, M. H. (1979). Best test design: Rasch measurement. Chicago: MESA Press.

Table 1

Sample Description

Native Language	N	Residence in English-Speaking Country	N	Amount of English Language Study	N
Arabic	3	0-6 Months	85	0-6 Months	9
Bahasa Malaysian	1	6 Months to 1 Year	15	6 Months to 1 Year	7
Chinese	9	1-2 Years	12	1-2 Years	18
French	11	2-3 Years	1	2-3 Years	11
German	17	3-5 Years	3	3-5 Years	15
Indonesian	2	More Than 5 Years	3	More than 5 Years	59
Italian	9	Not Reported	1	Not Reported	1
Japanese	30				
Korean	8				
Persian	3				
Polish	2				
Portuguese	2				
Spanish	13				
Swahili	1				
Thai	6				
Turkish	1				
Not Reported	2				
Totals	120		120		120

Table 2

Design of Listening Comprehension Test Administration

Listening Test 1	Listening Test 2	Listening Test 3
Passages:		
48 1-Sentence Passages	24 2-Sentence Passages	16 3-Sentence Passages
Items:		
48 Items (1 per Passage)	48 Items (2 per Passage)	48 Items (3 per Passage)
Repetition:		
First 24 Passages Repeated Once	First 12 Passages Repeated Once	First 8 Passages Repeated Once
Second 24 Passages Not Repeated	Second 12 Passages Not Repeated	Second 8 Passages Not Repeated
Length of Response Option:		
24 Short-Option Items	24 Short-Option Items	24 Short-Option Items
24 Longer-Option Items	24 Longer-Option Items	24 Longer-Option Items
Level of Comprehension Hierarchy:		
48 First-Level Items	24 First-Level Items	16 First-Level Items
	24 Second-Level Items	16 Second-Level Items
		16 Third-Level Items

. Test sequence was counterbalanced across all subjects so that each test was encountered in each of three sequences (i.e., 1-2-3, 2-3-1, or 3-1-2) by the same number of persons.

. Option length was randomly stratified within test across repetition and hierarchy conditions so that the same number of short- and longer-option items occurred under each condition.

. Hierarchy was necessarily confounded with passage length as it was defined by the number of stimulus sentences on which each answer depended.

. Items at hierarchical levels were met sequentially (1, 1-2, or 1-2-3) after each stimulus passage in accordance with passage length.

. 16 Items appeared in all three tests, 24 items appeared in two of the three tests, and 48 items appeared in no more than one test, in a balanced manner so that each subject had 48 one-time-item encounters, and 48 two-time-item encounters, and 48 three-time-item encounters.

. Each item was coded for sequence and number of encounters across the three tests in every experimental condition to enable control for possible contamination by these influences.

. Every test and every item was encountered by every subject with equal time allowed to each subject to respond.

. Most of these design features are evident within the actual test forms presented in the Appendix.

Table 3

Descriptive Statistics for All Tests

Test/Subtest	N Items	N Subjects	Mean	SD	α	α(50)
A. Listening Comprehension 1	48	120	31.992	7.119	0.835	0.841
/Repeated Items	24	120	17.742	3.963	0.766	0.872
/Nonrepeated Items	24	120	14.250	3.882	0.694	0.825
/Short-Option Items	24	120	15.533	3.968	0.720	0.843
/Long-Option Items	24	120	16.458	3.804	0.729	0.849
B. Listening Comprehension 2	48	120	31.367	8.114	0.871	0.875
/Repeated Items	24	120	15.508	4.300	0.780	0.881
/Nonrepeated Items	24	120	15.858	4.371	0.769	0.874
/Short-Option Items	24	120	16.192	4.458	0.783	0.883
/Long-Option Items	24	120	15.175	4.124	0.753	0.864
/1st Hierarchy Items	24	120	17.142	4.684	0.818	0.904
/2nd Hierarchy Items	24	120	14.225	3.296	0.708	0.835
C. Listening Comprehension 3	48	120	29.483	8.579	0.890	0.894
/Repeated Items	24	120	14.658	4.227	0.789	0.886
/Nonrepeated Items	24	120	14.825	4.879	0.834	0.913
/Short-Option Items	24	120	16.225	4.905	0.836	0.914
/Long-Option Items	24	120	13.258	4.173	0.773	0.876
/1st Hierarchy Items	16	120	11.075	3.325	0.769	0.912
/2nd Hierarchy Items	16	120	9.217	2.853	0.670	0.864
/3rd Hierarchy Items	16	120	9.192	3.234	0.729	0.894
D. TOEFL Reading Comprehension	30	120	17.942	5.430	0.781	0.859
E. Digital Memory	15	120	5.667	2.194	0.472	0.749

Table 4

Descriptive Statistics for Item Variances

(N = 144 Items)

Variable	Mean	SD	SE	Range
Difficulty:				
Proportion Correct (p)	.645	.192	---	.183 to .983
Rasch Difficulty (Subtest)	.000	1.086	.227	-3.380 to 2.510
Discrimination:				
Item-Total Point Biserial r (Subtest)	.379	.146	.091	- .108 to .648
Item-Total Point Biserial r (Total)	.332	.132	.091	- .076 to .575
Item-Total Biserial r (Subtest)	.526	.239	.092	- .312 to .916
Item-Total Biserial r (Total)	.458	.195	.092	- .220 to .746
Item-Total Biserial r (Subtest) (Part-Whole Overlap Corrected)	.484	.238	.092	- .328 to .907
Item-Total Biserial r (Total) (Part-Whole Overlap Corrected)	.412	.195	.092	- .238 to .723
Construct Validity:				
Rasch Model Infit (Subtest)	-.018	1.076	---	-3.000 to 2.840
TOEFL Reading Biserial r	.286	.147	.092	- .040 to .648
Digital Memory Biserial r	.170	.135	.092	- .101 to .452

. All correlation statistics employed Fisher Z transformations.

. Subtest estimates were based on the respective 48-item listening comprehension tests separately.

. Total test estimates were based on the composite of the three 48-item listening comprehension tests.

Table 5

Correlations Among Selected Item Variables
(N = 144 Items)

	REP	LEN	HIR	DIF	ITB	FIT	RDG	MEM
Repetition	1.000							
Option Length	-0.028	1.000						
Hierarchy	-0.000	0.040	1.000					
Rasch Difficulty	0.110	0.102	0.174	1.000				
Item-Total Biserial	-0.021	-0.122	-0.036	-0.185*	1.000			
Rasch Infit	0.130	0.121	0.132	0.166	-0.763**	1.000		
TOEFL Reading	-0.167	-0.044	-0.063	0.026	0.433**	-0.480**	1.000	
Digital Memory	-0.083	0.046	-0.078	-0.038	0.381**	-0.361**	0.397**	1.000

. Fisher Z transformations were employed in computations with correlation coefficients.

. Item-total biserials were based on the individual 48-item listening tests and were corrected for part-whole overlap.

*p < 0.05
**p < 0.01

Table 6A

ANOVA with Rasch Difficulty as Dependent Variable (N = 144 Items)

Source	DF	MS	F	P
Repetition	1	3.719	3.436	0.066
Option Length	1	6.265	5.787	0.018*
Hierarchy	2	1.199	1.107	0.334
REPxLEN	1	3.635	3.358	0.069
REPxHIR	2	0.010	0.009	0.991
LENxHIR	2	4.657	4.302	0.015*
REPxLENxHIR	2	1.762	1.762	0.176
Error	132	1.083		
(R = 0.391)				

*p < 0.05

Table 6B

ANCOVA with Rasch Difficulty as Dependent Variable and with Sequence and Encounters as Concomitant Variables (N = 144 Items)

Source	DF	MS	F	P
Repetition	1	4.121	3.861	0.052
Option Length	1	6.902	6.467	0.012*
Hierarchy	2	2.907	2.724	0.069
Sequence	1	3.451	3.233	0.074
Encounters	1	0.893	0.837	0.362
REPxLEN	1	3.102	2.906	0.091
REPxHIR	2	0.081	0.076	0.927
LENxHIR	2	4.964	4.652	0.011*
REPxLENxHIR	2	2.321	2.174	0.118
Error	130	1.067		
(R = 0.422)				

*p < 0.05

Table 6C

Means and Standard Deviations of Rasch Item Difficulty Estimates at All Levels of Significant Effects (N = 144 Items)

Main Effect	Level 1 N	Level 1 Mean	Level 1 SD	Level 2 N	Level 2 Mean	Level 2 SD	Level 3 N	Level 3 Mean	Level 3 SD
Repetition	72	-0.120	1.201	72	0.119	0.951	---	----	----
Option Length	72	-0.111	0.914	72	0.111	1.232	---	----	----
Hierarchy	88	-0.172	0.974	40	0.290	1.263	16	0.219	1.076
LENxHIR (L1)	46	-0.107	0.895	18	-0.054	1.072	8	-0.263	0.700
(L2)	42	-0.243	1.059	22	0.571	1.359	8	0.701	1.208

Table 7A

ANOVA with Overlap-Corrected Biserial Item-Total Individual Test Discrimination as Dependent Variable (N = 144 Items)

Source	DF	MS	F	P
Repetition	1	0.003	0.047	0.828
Option Length	1	0.147	2.405	0.123
Hierarchy	2	0.053	0.872	0.421
REPxLEN	1	0.002	0.025	0.873
REPxHIR	2	0.008	0.126	0.882
LENxHIR	2	0.038	0.614	0.543
REPxLENxHIR	2	0.018	0.295	0.745
Error	132	0.061		
(R = 0.215)				

Table 7B

ANCOVA with Overlap-Corrected Biserial Item-Total Individual Test Discrimination as Dependent Variable and with Sequence and Encounters as Concomitant Variables (N = 144 Items)

Source	DF	MS	F	P
Repetition	1	0.022	0.387	0.535
Option Length	1	0.185	3.225	0.075
Hierarchy	2	0.196	3.404	0.036*
Sequence	1	0.531	9.239	0.003**
Encounters	1	0.102	1.766	0.186
REPxLEN	1	0.000	0.004	0.951
REPxHIR	2	0.003	0.059	0.943
LENxHIR	2	0.046	0.796	0.453
REPxLENxHIR	2	0.007	0.127	0.881
Error	130	0.057		
(R = 0.344)				

*$p < 0.05$

**$p < 0.01$

Table 8A

ANOVA with Rasch Model FIT as Dependent Variable
(N = 144 Items)

Source	DF	MS	F	P
Repetition	1	2.156	1.946	0.165
Option Length	1	4.746	4.284	0.040*
Hierarchy	2	3.517	3.174	0.045*
REPxLEN	1	0.001	0.001	0.971
REPxHIR	2	0.251	0.227	0.797
LENxHIR	2	2.410	2.175	0.118
REPxLENxHIR	2	0.629	0.568	0.568
Error	132	1.108		
(R = 0.342)				

*p < 0.05

Table 8B

ANCOVA with Rasch Model Fit as Dependent Variable and with
Sequence and Encounters as Concomitant Variables (N = 144 Items)

Source	DF	MS	F	P
Repetition	1	1.858	1.672	0.198
Option Length	1	5.188	4.669	0.033*
Hierarchy	2	4.251	3.825	0.024*
Sequence	1	0.901	0.811	0.370
Encounters	1	0.982	0.884	0.349
REPxLEN	1	0.014	0.013	0.910
REPxHIR	2	0.180	0.162	0.851
LENxHIR	2	2.440	2.195	0.115
REPxLENxHIR	2	0.537	0.484	0.618
Error	130	1.111		
(R = 0.357)				

*p < 0.05

Table 9

Multitrait-Multimethod Validation Matrix with Two Levels Each of Repetition, Length and Hierarchy in Place of Traits and High/Low Difficulty Items in Place of Methods (N = 120 Subjects)

		REP1		REP2		LEN1		LEN2		HIR1		HIR2	
		H	L	H	L	H	L	H	L	H	L	H	L
REP1	H	1.000											
	L	0.437	1.000										
REP2	H	0.625	0.725	1.000									
	L	0.505	0.648	0.547	1.000								
LEN1	H	0.556	0.429	0.512	0.442	1.000							
	L	0.485	0.432	0.486	0.344	0.735	1.000						
LEN2	H	0.558	0.337	0.484	0.392	0.656	0.569	1.000					
	L	0.482	0.519	0.545	0.372	0.621	0.697	0.505	1.000				
HIR1	H	0.603	0.421	0.495	0.494	0.413	0.415	0.470	0.494	1.000			
	L	0.432	0.462	0.439	0.439	0.490	0.479	0.389	0.551	0.723	1.000		
HIR2	H	0.494	0.467	0.506	0.506	0.507	0.488	0.579	0.559	0.595	0.624	1.000	
	L	0.387	0.402	0.420	0.420	0.445	0.534	0.399	0.574	0.633	0.723	0.587	1.000
Mean		7.150	8.467	5.492	8.758	6.717	9.508	4.217	9.042	7.350	9.792	4.917	9.308
SD		2.441	2.173	2.308	2.086	3.202	2.021	2.294	2.495	2.659	2.365	2.266	2.122
N Items		12	12	12	12	12	12	12	12	12	12	12	12

APPENDIX

(INSTRUMENTS)

QUESTIONNAIRE

NAME: _____

NATIVE LANGUAGE: _____

How long have you lived in the United States or in any other English-speaking country?

_____ 0 - 6 months _____ 6 months to 1 year

_____ 1 - 2 years _____ 2 - 3 years

_____ 3 - 5 years _____ More than 5 years

How long have you studied English?

_____ 0 - 6 months _____ 6 months to 1 year

_____ 1 - 2 years _____ 2 - 3 years

_____ 3 - 5 years _____ More than 5 years

SCRIPT - 1

(Allow 12 seconds between items. Repeat 1-24. Do not repeat 25-48.)

1. You don't have to tell me if you don't feel like it.
2. Jane was asked to take one of the parts in the school play.
3. Whatever the consequences, I'm ready to try it.
4. Cindy had the shoemaker sharpen her ice skates.
5. He placed his chair so that he could see out the window.
6. The gas tank is empty.
7. Across the street is a park where we can eat our lunch.
8. We hardly studied at all last weekend.
9. Angela hopes to attend business school in the fall.
10. Why don't we move the chairs inside?
11. How boring this homework is!
12. He himself didn't know what to do.
13. His art was appreciated by the younger people at the exhibit.
14. Sam measured the flour, sugar, and spices and then mixed in the eggs.
15. He says he told the truth, but I don't believe him.
16. He doesn't teach in <u>this</u> department.
17. Mr. Hubbard served as chairman of the department until his retirement last year.
18. I found that poem hard to understand, didn't you?
19. I'll have to take this coat to the dry cleaner.
20. If your plane reservations aren't confirmed forty-eight hours in advance, they may be canceled.
21. I'm going to help Theresa with her math this afternoon.
22. Julie had better go to the supermarket right away because her sister is coming for lunch.
23. It seems as though we've known each other for a long time instead of just two weeks.

24. The motorcycle costs too much, don't you agree?

25. The outdoor concert was called off due to the weather.

26. In the basement I've discovered a defective heating unit that needs fixing.

27. Kate was really feeling down in the dumps about her latest chemistry assignment.

28. The person to see about housing is the dean of students.

29. There ought to be more pencils than those left in the box.

30. Can you read the signpost from here?

31. After the speech came a brief question-and-answer session.

32. She's been through a lot lately.

33. You can expect to spend at least an hour on this reading assignment.

34. Only Toby went to the movie.

35. Sue swims a mile every day to keep in shape.

36. Jeremy does his homework in the library with Sue.

37. That isn't all I want.

38. I wish I had photocopied that article so that I could refer to it now.

39. I bought this coat when I was abroad.

40. This trip'll be shorter on the subway than on the bus.

41. This television program is not in the least boring.

42. By the time we get to the airport, the plane will have taken off.

43. Whoever wins this game gets to play against Molly in the finals.

44. To accuse him of all people!

45. Nobody likes grapes more than I do.

46. The high winds resulted in heavy damage to trees and power lines.

47. Only Bill could draw a sketch like that.

48. Dick's parents made him spend his vacation at home.

NAME: _____

LISTENING COMPREHENSION - 1

Directions: For each question in this part you will hear a short sentence. Each sentence will be spoken one or two times. The sentences you hear will not be written out for you. Therefore, you must listen carefully to understand what the speaker says.

After you hear a sentence, read the four choices in your test book, marked (A), (B), (C), and (D), and decide which one is closest in meaning to the sentence you heard. Then, mark your answer on the test paper.

Example I

You will hear:

You will read: (A) Mary outswam the others.
(B) Mary ought to swim with them.
(C) Mary and her friends swam to the island.
(D) Mary's friends owned the island.

The speaker said, "Mary swam out to the island with her friends." Sentence (C), "Mary and her friends swam to the island," is closest in meaning to the sentence you heard. Therefore, you should choose answer (C).

Example II

You will hear:

You will read: (A) Please remind me to read this book.
(B) Could you help me carry these books?
(C) I don't mind if you help me.
(D) Do you have a heavy course load this term?

The speaker said, "Would you mind helping me with this load of books?" Sentence (B), "Could you help me carry these books?" is the closest in meaning to the sentence you heard. Therefore, you should choose answer (B).

1. You...
   (A) might tell me.
   (B) should tell me.
   (C) shouldn't tell me.
   (D) needn't tell me.

2. (A) Jane asked if she could be in the school play.
   (B) Jane took part of my lunch today.
   (C) Jane is very involved in her schoolwork.
   (D) Jane was offered a role in the play.

3. The consequences...
   (A) are already known.
   (B) won't stop me.
   (C) are known by trial.
   (D) won't ever change.

4. The shoemaker...
   (A) thought Cindy was nice.
   (B) sharpened Cindy's skates.
   (C) shined Cindy's shoes.
   (D) gave Cindy rice cakes.

5. He moved...
   (A) in his chair.
   (B) from the window.
   (C) to see better.
   (D) under his place.

6. (A) The tank is broken.
   (B) There's no gas left.
   (C) The gas is no good.
   (D) Thanks for the gas.

7. Let's...
   (A) part across the street.
   (B) eat in the park.
   (C) pack a lunch.
   (D) cross on a hunch.

8. We studied...
   (A) all last weekend.
   (B) at last.
   (C) all the material.
   (D) very little.

9. (A) Angela wants to begin business school this autumn.
   (B) Until she fell, Angela had been planning to go to school.
   (C) Angela plans to attend to her business at school.
   (D) Attendance ac Angela's school has declined.

10. (A) Aren't the chairs inside?
   (B) We don't know which chairs to move.
   (C) I think we should take the chairs in.
   (D) Why do you want to move the chairs?

11. This homework is...
   (A) less boring.
   (B) interesting, isn't it?
   (C) not boring, is it?
   (D) very uninteresting.

12. He...
   (A) didn't know either.
   (B) knew what to do.
   (C) didn't do it.
   (D) told what he knew.

13. (A) No one appreciated his art.
(B) The artist did not care for the people at the exhibit.
(C) The artist enjoyed having only young people at the exhibit.
(D) The younger people liked his art.

14. (A) Sam was doing a chemistry experiment.
(B) Sam was baking.
(C) Sam became confused.
(D) Sam measured up to expectations.

15. (A) He told me not to believe it.
(B) He thinks I don't tell the truth.
(C) I think his story is false.
(D) I don't believe he lied.

16. He...
(A) doesn't teach well.
(B) teaches elsewhere.
(C) teaches history.
(D) doesn't like teaching.

17. (A) Mr. Hubbard served the chairman.
(B) Mr. Hubbard replaced the last chairman.
(C) Mr. Hubbard is no longer the chairman.
(D) Mr. Hubbard was manager of the apartment.

18. (A) It's a difficult poem, isn't it?
(B) Didn't you find the poem we were assigned to read?
(C) Wasn't it hard to stand there and recite that poem?
(D) You lost the poem, didn't you?

19. (A) It's the only coat I have.
(B) I can take this coat for you.
(C) The dry cleaner has my coat.
(D) My coat needs cleaning.

20. Plane reservations should be...
(A) made in advance.
(B) confirmed.
(C) canceled.
(D) canceled after 48 hours.

21. Theresa...
(A) helps with math.
(B) goes this afternoon.
(C) will get help in math.
(D) teaches in the afternoon.

22. (A) Julie had lunch at the grocery store.
(B) Julie needs to buy some food quickly.
(C) Julie must write to her sister immediately.
(D) Julie got a better mark on the test than her sister.

23. (A) We've been friends for a long time.
(B) We haven't seen each other in a while.
(C) We met only two weeks ago.
(D) We hardly know anything about each other.

24. (A) How much did you agree to pay for the motorcycle?
(B) Don't you think that the motorcycle is too expensive?
(C) I don't agree with you about the cost of the motorcycle.
(D) I think you should agree to buy the motorcycle.

25. The concert...
(A) was overdue.
(B) was called together.
(C) rained out.

26. The basement...
(A) has a detective unit.
(B) heater needs repair.
(C) is well covered.
(D) prices need fixing.

27. Kate...
(A) taught chemistry classes.
(B) dumped chemical wastes.
(C) was down an assignment.
(D) disliked her assignment.

28. The dean of students...
(A) has someone to see.
(B) houses students.
(C) sees to housing.
(D) prevents carousing.

29. Some pencils...
(A) fell in the box.
(B) must be missing.
(C) cost too much.
(D) ought to be left.

30. (A) Isn't the post office near here?
(B) Where's the letter to be signed?
(C) Isn't that a side street?
(D) Can you see what that sign says?

31. (A) Only the brief questions were answered.
(B) Someone was asked to give a speech.
(C) The speaker spent a short while answering questions.
(D) Someone arrived after the speech to ask some questions.

32. (A) She's done a great deal of traveling.
(B) There's a good reason she's late.
(C) She's just finished her share of the work.
(D) Things have been difficult for her recently.

33. You will...
(A) expect an assignment.
(B) spend much for this.
(C) read less than an hour.
(D) need an hour or more.

34. (A) Toby went to the movie alone.
(B) Toby has only gone to the movie.
(C) Toby went to one movie.
(D) If only Toby would got to the movie!

35. Sue...
(A) is mild mannered.
(B) smiles every day.
(C) keeps escaping.
(D) swims daily.

36. Jeremy...
(A) works at home.
(B) studies with Sue.
(C) left the library.
(D) doesn't know Sue.

37. (A) That's just part of what I want.
(B) I don't want that at all.
(C) I want this rather than that.
(D) Other people want that, but I don't.

38. I should have...
(A) studied photography.
(B) made a copy.
(C) returned the article.
(D) preferred photocopy.

39. (A) I took this coat abroad with me.
(B) This coat is too big for me now.
(C) I purchased this coat while out of the country.
(D) This coat is very broad in the shoulders.

40. (A) Taking the subway would get us there faster.
(B) The bus and subway take the same amount of time.
(C) Going by bus would take less time.
(D) The bus goes past the subway station.

41. This program is...
(A) endlessly boring.
(B) quite interesting.
(C) the least boring.
(D) the least interesting.

42. We'll ...
(A) leave the airport.
(B) miss the plane.
(C) take time off.
(D) get the fare.

43. The winner ...
(A) finally played.
(B) plays Molly.
(C) is finalized.
(D) is Molly.

44. (A) No one accused him.
(B) They didn't act on his accusation.
(C) He accused all of them.
(D) I can't believe they accused him.

45. (A) I like grapes better than anyone does.
(B) I grow more grapes than anyone else.
(C) Grapes are more nutritious than I thought.
(D) Very few people like grapes.

46. (A) The strong winds broke tree limbs.
(B) The high winds and heavy seas made us feel helpless.
(C) The wind caused heavy flooding and drainage problems.
(D) The power lines damaged some big trees.

47. (A) No one else could draw such a picture.
(B) No one else was allowed to sketch.
(C) Bill drew only one sketch.
(D) Bill draws only what he likes.

48. Dick...
(A) loved his parents.
(B) spent his money.
(C) vacationed at home.
(D) left his home.

SCRIPT - 2

(Allow 20 seconds between items. Repeat passages 1-12. Do not repeat passages 13-24.)

Items

1 to 2 — 1. You don't have to tell me if you don't feel like it.
You're welcome to keep it a secret if you wish.

3 to 4 — 2. Jane was asked to take one of the parts in the school play.
She has performed well in school productions since she was a child.

5 to 6 — 3. Whatever the consequences, I'm ready to try it.
It sounds like an exciting thing to do.

7 to 8 — 4. Cindy had the shoemaker sharpen her ice skates.
She was getting ready for the race.

9 to 10 — 5. He placed his chair so that he could see out the window.
He always enjoyed the view of the valley in the springtime.

11 to 12 — 6. The gas tank is empty.
You'd better stop soon.

13 to 14 — 7. Across the street is a park where we can eat our lunch.
There are lots of picnic tables and it is usually quiet.

15 to 16 — 8. We hardly studied at all last weekend.
Our family came for a short visit.

17 to 18 — 9. Angela hopes to attend business school in the fall.
She believes she wants to be an accountant.

19 to 20 — 10. Why don't we move the chairs inside?
It's cold here when the wind blows.

21 to 22 — 11. How boring this homework is!
Anything else is more interesting.

23 to 24 — 12. He himself didn't know what to do.
But he pretended to know the answers.

25 to 26 — 13. His art was appreciated by the younger people at the exhibit.
But the older people were sure that he had no talent.

27 to 28 — 14. Sam measured the flour, sugar, and spices and then mixed in the eggs.
Then he stirred it for several minutes and put it in the oven.

29 to 30 — 15. He says he told the truth, but I don't believe him.
He has a history of stretching the facts to suit himself.

31 to 32 16. He doesn't teach in <u>this</u> department.
Maybe you should ask next door.

33 to 34 17. Mr. Hubbard served as chairman of the department until his retirement last year.
His efforts to hire outstanding new faculty members will always be appreciated.

35 to 36 18. I found that poem hard to understand, didn't you?
I couldn't even decide what the main idea was.

37 to 38 19. I'll have to take this coat to the dry cleaner.
It's got food stains on the collar and both sleeves.

39 to 40 20. If your plane reservations aren't confirmed forty-eight hours in advance, they may be canceled.
So you'd better get on the phone in the office across the hall.

41 to 42 21. I'm going to help Theresa with her math this afternoon.
She's having trouble with long division and fractions.

43 to 44 22. Julie had better go to the supermarket right away because her sister is coming for lunch.
There's no milk or sandwich bread anywhere in the house

45 to 46 23. It seems as though we've known each other for a long time instead of just two weeks.
It must be because we have so much in common and agree about so many things.

47 to 48 24. The motorcycle costs too much, don't you agree?
We could find a car for that price.

NAME: _____

LISTENING COMPREHENSION - 2

Directions: For each question in this part you will hear a short passage. Each passage will be spoken one or two times. The sentences you hear will not be written out for you. Therefore, you must listen carefully to understand what the speaker says.

After you hear a passage, read the four choices, marked (A), (B), (C), and (D), for each question, and decide which one is closest in meaning to the passage you heard. Then, mark your answer on the test paper. Answer two questions after each passage.

1. You...
   (A) might tell me.
   (B) should tell me.
   (C) shouldn't tell me.
   (D) needn't tell me.

2. (A) You're thinking about telling me.
   (B) You like to keep secrets from me.
   (C) You don't feel welcome here.
   (D) You secretly wish to feel welcome.

* * * * *

3. (A) Jane asked if she could be in the school play.
   (B) Jane took part of my lunch today.
   (C) Jane is very involved in her schoolwork.
   (D) Jane was offered a role in the play.

4. (A) Jane is a child who only wants to play.
   (B) Jane always takes part in school.
   (C) Jane was a good choice since she does well.
   (D) To perform well takes much practice.

* * * * *

5. The consequences...
   (A) are already known.
   (B) won't stop me.
   (C) are known by trial.
   (D) won't ever change.

6. (A) I tried it already and liked it.
   (B) I like whatever happens to me.
   (C) The consequences will be very nice.
   (D) It is so much fun that it is worth any trouble.

* * * * *

7. The shoemaker...
   (A) thought Cindy was nice.
   (B) sharpened Cindy's skates.
   (C) shined Cindy's shoes.
   (D) gave Cindy rice cakes.

8. (A) Shoemakers are usually ready.
   (B) Racing requires sharp skates.
   (C) Cakes must be prepared.
   (D) Cindy fell on her face.

* * * * *

9. He moved...
   (A) in his chair.
   (B) from the window.
   (C) to see better.
   (D) under his place.

10. (A) He called a rally in the springtime.
    (B) The view could be enjoyed after moving.
    (C) He could see his chair from the window.
    (D) Springtime was a difficult time.

* * * * *

11. (A) The tank is broken.
    (B) There's no gas left.
    (C) The gas is no good.
    (D) Thanks for the gas.

12. (A) Stop emptying the gas from the tank.
    (B) It's better to stop giving thanks.
    (C) Empty the gas tank quickly.
    (D) You need to get gas right away.

* * * * *

13. Let's...
    (A) part across the street.
    (B) eat in the park.
    (C) pack a lunch.
    (D) cross on a hunch.

14. There's ...
    (A) a good place to eat.
    (B) some lunch in the park.
    (C) quite a stable.
    (D) parking in the street.

* * * * *

15. We studied...
    (A) all last weekend.
    (B) at last.
    (C) all the material.
    (D) very little.

16. (A) The weekend seemed entirely too short.
    (B) Our studies lasted all weekend.
    (C) A family visit interrupted our studies.
    (D) We studied hard for our family.

* * * * *

17. (A) Angela wants to begin business school this autumn.
    (B) Until she fell, Angela had been planning to go to school.
    (C) Angela plans to attend to her business at school.
    (D) Attendance at Angela's school has declined.

18. Angela wants ...
    (A) No business at all.
    (B) School to end.
    (C) An accounting course.
    (D) To count her school.

* * * * *

19. (A) Aren't the chairs inside?
    (B) We don't know which chairs to move.
    (C) I think we should take the chairs in.
    (D) Why do you want to move the chairs?

20. Let's ...
    (A) move out of the wind.
    (B) not move inside.
    (C) get new chairs.
    (D) chair that motion.

* * * * *

21. This homework is...
    (A) less boring.
    (B) interesting, isn't it?
    (C) not boring, is it?
    (D) very uninteresting.

22. (A) I'd prefer anything to this homework.
    (B) I'm sorry to be so boring.
    (C) My homework is more interesting.
    (D) I like work more than anything.

* * * * *

23. He...
    (A) didn't know either.
    (B) knew what to do.
    (C) didn't do it.
    (D) told what he knew.

24. (A) He tended to know every answer.
    (B) He gave a false impression of his knowledge.
    (C) He knew more than he was willing to show.
    (D) He didn't know where he had learned the answers.

* * * * *

25. (A) No one appreciated his art.
(B) The artist did not care for the people at the exhibit.
(C) The artist enjoyed having only young people at the exhibit.
(D) The younger people liked his art.

26. Not everyone ...
(A) likes art.
(B) has talent.
(C) appreciates young people.
(D) has the same taste.

* * * * *

27. (A) Sam was doing a chemistry experiment.
(B) Sam was baking.
(C) Sam became confused.
(D) Sam measured up to expectations.

28. (A) Sam seldom stirred when he was busy.
(B) Sam followed a plan in his cooking.
(C) Sam tried to bake one dozen.
(D) Sam was mixed up about what he was doing.

* * * * *

29. (A) He told me not to believe it.
(B) He thinks I don't tell the truth.
(C) I think his story is false.
(D) I don't believe he lied.

30. He ...
(A) is often untruthful.
(B) says that suit stretched.
(C) has an unbelievable history.
(D) told the truth.

* * * * *

31. He...
(A) doesn't teach well.
(B) teaches elsewhere.
(C) teaches history.
(D) doesn't like teaching.

32. He ...
(A) might teach nearby.
(B) asked the department.
(C) came next door.
(D) doesn't teach as asked.

* * * * *

33. (A) Mr. Hubbard served the chairman.
(B) Mr. Hubbard replaced the last chairman.
(C) Mr. Hubbard is no longer the chairman.
(D) Mr. Hubbard was manager of the apartment.

34. He will be ...
(A) chairman of the department.
(B) retiring next year.
(C) a new faculty member.
(D) remembered for recruitment.

* * * * *

35. (A) It's a difficult poem, isn't it?
(B) Didn't you find the poem we were assigned to read?
(C) Wasn't it hard to stand there and recite that poem?
(D) You lost the poem, didn't you?

36. I ...
(A) have an idea for a poem.
(B) decided, didn't you?
(C) had trouble understanding it.
(D) found that poem.

* * * * *

37. (A) It's the only coat I have.
(B) I can take this coat for you.
(C) The dry cleaner has my coat.
(D) My coat needs cleaning.

38. The cleaner will...
(A) remove stains.
(B) sew the sleeves.
(C) take great pains.
(D) press the collar.

* * * * *

39. Plane reservations should be...
(A) made in advance.
(B) confirmed.
(C) canceled.
(D) canceled after 48 hours.

40. (A) They confirm reservations in that office.
(B) There's someone on the phone in that office.
(C) Use the phone to confirm your plane.
(D) They canceled 48 planes by phone.

* * * * *

41. Theresa...
(A) helps with math.
(B) goes this afternoon.
(C) will get help in math.
(D) teaches in the afternoon.

42. Theresa needs...
(A) a fraction of your time.
(B) a new math book.
(C) my help with division.
(D) tomorrow afternoon.

* * * * *

43. (A) Julie had lunch at the grocery store.
(B) Julie needs to buy some food quickly.
(C) Julie must write to her sister immediately.
(D) Julie got a better mark on the test than her sister.

44. Julie should...
(A) get bread for her sister.
(B) come home for lunch.
(C) stay in the house.
(D) use up the milk.

* * * * *

45. (A) We've been friends for a long time.
(B) We haven't seen each other in a while.
(C) We met only two weeks ago.
(D) We hardly know anything about each other.

46. We...
(A) get along well with others.
(B) commonly disagree.
(C) have recently become friends.
(D) stayed a long time.

* * * * *

47. (A) How much did you agree to pay for the motorcycle?
(B) Don't you think that the motorcycle is too expensive?
(C) I don't agree with you about the cost of the motorcycle.
(D) I think you should agree to buy the motorcycle.

48. (A) We might as well buy a car.
(B) We don't agree about the price.
(C) Let's buy a better motorcycle.
(D) Cars cost too much money anyway.

* * * * *

Script - 3

(Allow 30 seconds between items. Repeat passages 1-8. Do not repeat passages 9-16.)

Items

1 to 3 1. You don't have to tell me if you don't feel like it.
You're welcome to keep it a secret if you wish.
But I have some news you might like to hear too.

4 to 6 2. Jane was asked to take one of the parts in the school play.
She has performed well in school productions since she was a child.
And I think she wants to be a high school drama teacher some day.

7 to 9 3. Whatever the consequences, I'm ready to try it.
It sounds like an exciting thing to do.
And we've had no entertainment for weeks.

10 to 12 4. Cindy had the shoemaker sharpen her ice skates.
She was getting ready for the race.
And she wanted to have every possible advantage.

13 to 15 5. He placed his chair so that he could see out the window.
He always enjoyed the view of the valley in the springtime.
It reminded him of his youth and the freedom he loved.

16 to 18 6. The gas tank is empty.
You'd better stop soon.
There's a station up ahead.

19 to 21 7. Across the street is a park where we can eat our lunch.
There are lots of picnic tables and it is usually quiet.
We can discuss that private matter and not be overheard.

22 to 24 8. We hardly studied at all last weekend.
Our family came for a short visit.
We went for a drive and talked.

25 to 27 9. Angela hopes to attend business school in the fall.
She believes she wants to be an accountant.
It will be an opportunity for her to use her math skills.

28 to 30 10. Why don't we move the chairs inside?
It's cold here when the wind blows.
And I'm still getting over a cold.

31 to 33 11. How boring this homework is!
Anything else is more interesting.
I can't wait til this term is over.

34 to 36 12. He himself didn't know what to do.
But he pretended to know the answers.
And chided us for responding slowly.

37 to 39 13. His art was appreciated by the younger people at the exhibit.
But the older people were sure that he had no talent.
No art form has ever been uniformly liked by everyone.

40 to 42 14. Sam measured the flour, sugar, and spices and then mixed in the eggs.
Then he stirred it for several minutes and put it in the oven.
I waited impatiently as the cake was baking, hoping for a taste.

43 to 45 15. He says he told the truth, but I don't believe him.
He has a history of stretching the facts to suit himself.
And I can never know whether he's serious or just kidding.

46 to 48 16. He doesn't teach in <u>this</u> department.
Maybe you should ask next door.
They have more teachers than we do.

NAME: _____

LISTENING COMPREHENSION - 3

Directions: For each question in this part you will hear a short passage. Each passage will be spoken one or two times. The sentences you hear will not be written out for you. Therefore, you must listen carefully to understand what the speaker says.

After you hear a passage, read the four choices, marked (A), (B), (C), and (D), for each question, and decide which one is closest in meaning to the passage you heard. Then, mark your answer on the test paper. Answer three questions after each passage.

1. You...
   (A) might tell me.
   (B) should tell me.
   (C) shouldn't tell me.
   (D) needn't tell me.

2. (A) You're thinking about telling me.
   (B) You like to keep secrets from me.
   (C) You don't feel welcome here.
   (D) You secretly wish to feel welcome.

3. (A) You'd better not tell the secret.
   (B) You're welcome to hear the news.
   (C) I feel happy about whatever you decide.
   (D) If you don't share, I won't either.

* * * * *

4. (A) Jane asked if she could be in the school play.
   (B) Jane took part of my lunch today.
   (C) Jane is very involved in her schoolwork.
   (D) Jane was offered a role in the play.

5. (A) Jane is a child who only wants to play.
   (B) Jane always takes part in school.
   (C) Jane was a good choice since she does well.
   (D) To perform well takes much practice.

6. Jane's participation...
   (A) was quite a production.
   (B) will help her career.
   (C) was child's play.
   (D) will cost a lot.

* * * * *

7. The consequences...
   (A) are already known.
   (B) won't stop me.
   (C) are known by trial.
   (D) won't ever change.

8. (A) I tried it already and liked it.
   (B) I like whatever happens to me.
   (C) The consequences will be very nice.
   (D) It is so much fun that it is worth any trouble.

9. (A) Our lack of amusement makes me try strange things.
   (B) We won't know the consequences for weeks.
   (C) We can't try this new experiment.
   (D) It's fun no matter how it sounds.

* * * * *

10. The shoemaker...
   (A) thought Cindy was nice.
   (B) sharpened Cindy's skates.
   (C) shined Cindy's shoes.
   (D) gave Cindy rice cakes.

11. (A) Shoemakers are usually ready.
   (B) Racing requires sharp skates.
   (C) Cakes must be prepared.
   (D) Cindy fell on her face.

12. (A) The shoemaker was ready for the race.
   (B) It's possible to have your cake and eat it too.
   (C) It was a racing advantage to have sharpened skates.
   (D) Cindy fell again at the same place.

* * * * *

13. He moved...
    (A) in his chair.
    (B) from the window.
    (C) to see better.
    (D) under his place.

14. (A) He called a rally in the springtime.
    (B) The view could be enjoyed after moving.
    (C) He could see his chair from the window.
    (D) Springtime was a difficult time.

15. He moved his chair...
    (A) To get out of the sun.
    (B) Because he fixed the springs.
    (C) To get out of view from the window.
    (D) For memories brought by the view.

* * * * *

16. (A) The tank is broken.
    (B) There's no gas left.
    (C) The gas is no good.
    (D) Thanks for the gas.

17. (A) Stop emptying the gas from the tank.
    (B) It's better to stop giving thanks.
    (C) Empty the gas tank quickly.
    (D) You need to get gas right away.

18. You should stop...
    (A) to empty the gas tank.
    (B) to say thanks for gas.
    (C) at the station for gas.
    (D) quickly for directions.

* * * * *

19. Let's...
    (A) part across the street.
    (B) eat in the park.
    (C) pack a lunch.
    (D) cross on a hunch.

20. There's ...
    (A) a good place to eat.
    (B) some lunch in the park.
    (C) quite a stable.
    (D) parking in the street.

21. In the park...
    (A) people leave their cars.
    (B) we can eat and talk alone.
    (C) it matters to be overheard.
    (D) our friends have a picnic.

* * * * *

22. We studied...
    (A) all last weekend.
    (B) at last.
    (C) all the material.
    (D) very little.

23. (A) The weekend seemed entirely too short.
    (B) Our studies lasted all weekend.
    (C) A family visit interrupted our studies.
    (D) We studied hard for our family.

24. (A) We studied all last weekend in spite of the visit.
    (B) We had a short visit to study hall.
    (C) Our family studied last weekend.
    (D) Driving and talking with family took time from study.

* * * * *

25. (A) Angela wants to begin business school this autumn.
(B) Until she fell, Angela had been planning to go to school.
(C) Angela plans to attend to her business at school.
(D) Attendance at Angela's school has declined.

26. Angela wants ...
(A) no business at all.
(B) school to end.
(C) an accounting course.
(D) to count her school.

27. (A) Angela hopes to finish her business in fall.
(B) Angela's math skill is suited to accounting study.
(C) Angela finished school last fall.
(D) Angela counted on no new opportunities.

* * * * *

28. (A) Aren't the chairs inside?
(B) We don't know which chairs to move.
(C) I think we should take the chairs in.
(D) Why do you want to move the chairs?

29. Let's ...
(A) move out of the wind.
(B) not move inside.
(C) get new chairs.
(D) chair that motion.

30. Moving the chairs...
(A) will keep us outside.
(B) is done by the wind.
(C) will help my cold.
(D) is too cold to do.

* * * * *

31. This homework is...
(A) less boring.
(B) interesting, isn't it?
(C) not boring, is it?
(D) very uninteresting.

32. (A) I'd prefer anything to this homework.
(B) I'm sorry to be so boring.
(C) My homework is more interesting.
(D) I like work more than anything.

33. My boredom...
(A) is unexplainable.
(B) is better than work.
(C) is from not enough homework.
(D) will end after this term.

* * * * *

34. He...
(A) didn't know either.
(B) knew what to do.
(C) didn't do it.
(D) told what he knew.

35. (A) He tended to know every answer.
(B) He gave a false impression of his knowledge.
(C) He knew more than he was willing to show.
(D) He didn't know where he had learned the answers.

36. (A) He slowly tended to learn what to do.
(B) His answers were two-sided.
(C) He criticized our responses, but he knew less.
(D) He didn't know how to respond slowly.

* * * * *

37. (A) No one appreciated his art.
(B) The artist did not care for the people at the exhibit.
(C) The artist enjoyed having only young people at the exhibit.
(D) The younger people liked his art.

38. Not everyone ...
(A) likes art.
(B) has talent.
(C) appreciates young people.
(D) has the same taste.

39. Art forms...
(A) are not liked by youth.
(B) are sure to require talent.
(C) are not equally appreciated.
(D) are exhibited by older people.

* * * * *

40. (A) Sam was doing a chemistry experiment.
(B) Sam was baking.
(C) Sam became confused.
(D) Sam measured up to expectations.

41. (A) Sam seldom stirred when he was busy.
(B) Sam followed a plan in his cooking.
(C) Sam tried to bake one dozen.
(D) Sam was mixed up about what he was doing.

42. Sam's work...
(A) made me hungry.
(B) mixed me up.
(C) tired me out.
(D) stirred us all.

* * * * *

43. (A) He told me not to believe it.
(B) He thinks I don't tell the truth.
(C) I think his story is false.
(D) I don't believe he lied.

44. He ...
(A) is often untruthful.
(B) says that suit stretched.
(C) has an unbelievable history.
(D) told the truth.

45. (A) His manner of speaking makes him hard to believe.
(B) He is usually far too serious.
(C) He tells the truth no matter what happens.
(D) His kidding is a source of enjoyment.

* * * * *

46. He...
(A) doesn't teach well.
(B) teaches elsewhere.
(C) teaches history.
(D) doesn't like teaching.

47. He ...
(A) might teach nearby.
(B) asked the department.
(C) came next door.
(D) doesn't teach as asked.

48. (A) He has already finished his teaching.
(B) This department has offices next door.
(C) Since he doesn't work here, perhaps he is in a larger department.
(D) Maybe he is not a teacher after all.

* * * * *

3 3 3 3 3 3 3

Directions In the rest of this section you will read several passages. Each one is followed by several questions about it. For questions 31-60, you are to choose the one best answer, (A), (B), (C), or (D), to each question. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen.

Answer all questions following a passage on the basis of what is stated or implied in that passage.

Read the following passage:

> The rattles with which a rattlesnake warns of its presence are formed by loosely interlocking hollow rings of hard skin, which make a buzzing sound when its tail is shaken. As a baby, the snake begins to form its rattles from the button at the very tip of its tail. Thereafter, each time it sheds its skin, a new ring is formed. Popular belief holds that a snake's age can be told by counting the rings, but this idea is fallacious. In fact, a snake may lose its old skin as often as four times a year. Also, rattles tend to wear or break off with time.

Example I — Sample Answer: Ⓐ ● Ⓒ Ⓓ

A rattlesnake's rattles are made of

(A) skin
(B) bone
(C) wood
(D) muscle

According to the passage, a rattlesnake's rattles are made out of rings of hard skin. Therefore, you should choose answer (A).

Example II — Sample Answer: Ⓐ Ⓑ ● Ⓓ

How often does a rattlesnake shed its skin?

(A) Once every four years
(B) Once every four months
(C) Up to four times every year
(D) Four times more often than other snakes

The passage states that "a snake may lose its old skin as often as four times a year." Therefore, you should choose answer (C).

After you read the directions, begin work on the questions.

GO ON TO THE NEXT PAGE

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 31-35

As the result of several discoveries and the development of new techniques and equipment, the importance of the shellfish industry has increased since 1950. One shellfish, the oyster, has been cultivated in great numbers by means of the new technology. Marine biologists discovered that by using chemicals and raising the water temperature, they could induce oysters to lay eggs not only in the summer but also in the fall, winter, and spring. They also succeeded in breeding new strains of oysters that were resistant to diseases and grew faster and larger. In addition, the cultivated oysters taste better!

31. This passage was most likely taken from a book on
    (A) mixing chemicals
    (B) raising shellfish
    (C) building factories
    (D) treating diseases

32. The main topic of this passage is
    (A) an explanation of marine biology
    (B) new locations for oyster beds
    (C) improvements in the oyster industry
    (D) the shellfish industry before 1950

33. It can be inferred from the passage that before the use of the new technology, oysters laid their eggs each year during
    (A) one season
    (B) two seasons
    (C) three seasons
    (D) four seasons

34. According to the passage, all of the following are characteristics of the new strains of oysters EXCEPT
    (A) freedom from illness
    (B) more rapid growth
    (C) a larger size
    (D) flawless pearls

35. According to the passage, an unexpected result of the application of the technology was the oysters'
    (A) improved taste
    (B) increased yield
    (C) larger eggs
    (D) faster movement

GO ON TO THE NEXT PAGE

BEST COPY AVAILABLE

Questions 36-43

The mistletoe plant grows on deciduous trees—those that lose their leaves in autumn. It is spread by birds as they eat its white berries and then wipe their beaks on the bark of nearby branches. The sticky seeds put out tiny rootlets, thrust up leaves, and begin an extremely long life.

Insects leave mistletoe alone. Winds never blow it down. Wintry ice and parching summers cause it no harm. Virtually indestructible, it dies only when the tree dies. One mistletoe ball was estimated to have survived four hundred years. For the tree, it is only a minor pest, manufacturing its own food from the chlorophyll of its leaves and using the tree simply as a source of liquid and vital minerals.

36. The passage indicates that mistletoe can be found on trees that

(A) have sticky bark
(B) shed their leaves annually
(C) have exposed roots
(D) produce white berries

37. According to the passage, mistletoe seeds travel from place to place by

(A) clinging to birds' beaks
(B) sticking to berries
(C) spreading over loose bark
(D) blowing to nearby branches

38. According to the passage, mistletoe grows by

(A) attaching itself to leaves
(B) stretching out in long vines
(C) slowly climbing tree trunks
(D) sprouting roots into tree bark

39. The passage indicates that mistletoe plants are most likely to be damaged by

(A) bugs and animals
(B) extreme heat and cold
(C) destruction of trees
(D) disturbance by wind

GO ON TO THE NEXT PAGE

40. According to the passage, mistletoe may live

(A) until dry weather
(B) as long as the tree has its leaves
(C) an unusually long time
(D) a year at most

41. According to the passage, mistletoe obtains part of its nourishment by

(A) producing it independently
(B) removing it from tree leaves
(C) absorbing it from moist air
(D) utilizing dirt that sticks to bark

42. According to the passage, mistletoe does which of the following when it is growing on a tree?

(A) Robs the tree of chlorophyll.
(B) Attracts many insect pests.
(C) Destroys green leaves.
(D) Takes fluids from the tree.

43. Which of the following would be the best title for the passage?

(A) A Plant Pest
(B) A Durable Plant
(C) An Independent Plant
(D) A Plant with Many Uses

GO ON TO THE NEXT PAGE

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 44-49

The economy of the United States after 1952 was the economy of a well-fed, almost fully employed people. Despite occasional alarms, the country escaped any postwar liquidation and lived in a state of boom. The history of extraction, production, and distribution had therefore been almost nothing but a statistical table reflecting prosperity. An economic survey of the year 1955, a typical year of the 1950's, may be illuminating as illustrating the decade. The national output was valued at about 10 percent above that of 1954 (1955 output was estimated at 392 billion dollars). The production of manufacturers was about 40 percent more than it had averaged in the years immediately following World War II. The country's businesses spent about 30 billion dollars for new factories and machinery. National income available for spending was almost a third greater than it had been in 1950. Consumers spent about 256 billion dollars, that is, about 700 million dollars a day, or about twenty-five million dollars every hour, all around the clock. Sixty-five million people held jobs and only a little more than two million wanted jobs but could not find them. Only agriculture complained that it was not sharing in the boom. To some observers this was an ominous echo of the mid-1920's. As farmers' share of their products declined, marketing costs rose. But there were few pessimists among the observers of the national economy. Those few seemed to fear that the prosperity was based on government pump-priming on a stupendous scale.

44. The passage is mainly concerned with which of the following aspects of United States history?

(A) The agricultural trends of the 1950's
(B) The unemployment rate in 1955
(C) The general economic situation in the 1950's
(D) The federal budget of 1952

45. In the first paragraph, the word "boom" could best be replaced by which of the following?

(A) nearby explosion
(B) thunderous noise
(C) general public support
(D) rapid economic growth

46. The passage states that income available for spending in the United States was approximately how much greater in 1955 than in 1950?

(A) 33%
(B) 50%
(C) 60%
(D) 90%

47. According to the passage, about how many million people were unemployed in 1955?

(A) One
(B) Two
(C) Three
(D) Four

48. It can be inferred from the passage that most people in the United States in 1955 viewed the national economy with an air of

(A) optimism
(B) confusion
(C) derision
(D) suspicion

49. According to the passage, which of the following were LEAST satisfied with the national economy in the 1950's?

(A) Farmers
(B) Economists
(C) Politicians
(D) Steelworkers

GO ON TO THE NEXT PAGE

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 50-53

One of the extraordinary natural phenomena that disappeared as the size of American forests decreased was the occasional mass emigration of the gray squirrel. Caused by an excessive build-up of their populations, these movements involved millions of the animals—a half-billion were estimated to have migrated across southern Wisconsin in 1843. Moving in hordes, they devoured crops and even managed to struggle across such mighty rivers as the Ohio, thousands drowning in the effort. Currently, the squirrel population, although smaller, is on the increase; emigrations still occur but they are not often noticeable.

50. Gray squirrels usually emigrate because of

(A) major weather changes
(B) an increase in the number of squirrels
(C) humans building houses in nesting areas
(D) danger from hunters

51. Why do gray squirrels emigrate less frequently now than they used to?

(A) Most of them have already moved south.
(B) Most of them are impeded by the current of the Ohio River.
(C) There are now fewer natural enemies in the squirrels' habitats.
(D) There is now less forest land in which squirrels are concentrated.

52. According to the passage, the squirrel emigrations have caused problems for people because of

(A) damage to homes
(B) loss of farm produce
(C) disruption of river traffic
(D) harm to children

53. What is happening to the squirrel population today?

(A) It is gradually getting larger.
(B) It is being closely observed by biologists.
(C) It is becoming geographically more widespread.
(D) It is producing young that are smaller in size.

GO ON TO THE NEXT PAGE

BEST COPY AVAILABLE

Questions 54-60

The acknowledged "King of Ragtime" was the Black pianist and composer Scott Joplin. Joplin (1868-1917), originally from Texarkana, Texas, began his career as an itinerant pianist. By 1885 he was in St. Louis, playing in honky-tonks and sporting houses. He went to Chicago briefly (1893) to try his luck in the entertainment halls that had sprung up around the World's Fair, then in 1894 to Sedalia, Missouri, to stay until the turn of the century. His first published rag, *Original Rags*, came out in March, 1899; later the same year appeared *Maple Leaf Rag*, named for a saloon and dance hall in Sedalia. The work was an instant and resounding success, and by the time of his death Joplin had published more than thirty original rags, and other piano pieces, songs, and arrangements. He had even larger aims: in 1902 he finished a ballet score called *Rag Time Dance*, and in 1903 the opera *A Guest of Honor*, unpublished and now apparently lost; in 1911 came another opera, *Treemonisha*. The artistic success of these larger works is debatable, but that of Joplin's piano rags is not; they can only be described as elegant, varied, often subtle, and as sharply incised as a cameo. They are the precise American equivalent, in terms of a native style of dance music, of minuets by Mozart, mazurkas by Chopin, or waltzes by Brahms. They can be both lovely and powerful, infectious and moving—depending, of course, on the skill and stylishness of the pianist; for they are not easy music technically and they demand a clean but "swinging" performance.

54. Which of the following is the best title for the passage?

(A) Scott Joplin's Early Career
(B) Rare Piano Works of Scott Joplin
(C) Sedalia: The Birthplace of Ragtime
(D) A Ragtime Composer and His Music

55. It can be inferred from the passage that Joplin is recognized as the "King of Ragtime" because he

(A) was probably the greatest composer of ragtime music
(B) began his career as a famous child pianist
(C) created the character "King of Ragtime" in one of his operas
(D) was a descendant of a European royal family

56. Where was Scott Joplin born?

(A) Sedalia
(B) St. Louis
(C) Texarkana
(D) Chicago

57. According to the passage, which of the following is an accurate statement about *Maple Leaf Rag*?

(A) It was Joplin's favorite composition.
(B) Its name came from an establishment in Missouri.
(C) It was published in March 1899.
(D) Its popularity grew slowly.

GO ON TO THE NEXT PAGE

58. The phrase "these larger works" in the second half of the passage refers to Joplin's

(A) piano rags
(B) operas and ballet
(C) arrangements of other composers' music
(D) music published before *Original Rags*

59. Toward the end of the passage, the author refers to the works of other composers in order to illustrate the

(A) popularity of different styles of dance music in recent centuries
(B) success of Joplin's operas in Europe
(C) high quality of Joplin's work as an American musical form
(D) powerful movement attributed to Joplin's compositions

60. From the last sentence of the passage, one may infer that Joplin's piano music can best be appreciated when played

(A) by a highly skilled pianist
(B) in an elegant setting
(C) with a moving, classical style
(D) for a small audience

THIS IS THE END OF SECTION 3

IF YOU FINISH BEFORE TIME IS CALLED, CHECK YOUR WORK ON SECTION 3 ONLY.
DO NOT READ OR WORK ON ANY OTHER SECTION OF THE TEST.

STOP STOP STOP STOP STOP STOP STOP

BEST COPY AVAILABLE

DIGITAL MEMORY TEST

DIRECTIONS (presented orally by recording): Listen carefully to the following numbers. Try to remember as many numbers as you can. Do not write anything until you are told. After listening, you will have two minutes to write as many of these numbers as you can remember. Now listen carefully.

4
39
5
44
14
84
1
8
43
92
29
12
59
48
50

Now write as many numbers as you remember.

BEST COPY AVAILABLE

Cover Printed on Recycled Paper
57906-01194 • Y31M 5 • 275700 • Printed in U.S.A.